ESTIMATION OF PV LOCATION IN DISTRIBUTION SYSTEMS BASED ON VOLTAGE SENSITIVITIES

A Dissertation
Presented to
The Academic Faculty

by

Cristian Gómez Peces

In Partial Fulfillment
of the Requirements for the Degree
Master of Science in Electrical and Computer Engineering in the
School of Electrical and Computer Engineering

Georgia Institute of Technology
December 2020

ESTIMATION OF PV LOCATION IN DISTRIBUTION SYSTEMS BASED ON VOLTAGE SENSITIVITIES

Approved by:

Dr. Santiago Grijalva, Advisor
School of Electrical and Computer Engineering
Georgia Institute of Technology

Dr. Sakis Meliopoulos
School of Electrical and Computer Engineering
Georgia Institute of Technology

Dr. Lukas Graber
School of Electrical and Computer Engineering
Georgia Institute of Technology

Date Approved: December 4, 2020

ACKNOWLEDGEMENTS

I would like to thank my advisor, Dr. Santiago Grijalva, for his guidance and invaluable feedback throughout this thesis. From day one, he provided me with full freedom to conduct research, and gave me excellent mentorship in how to properly present ideas, and how to structure approaches to complex problems. It has been a great experience working alongside him to bring these solutions to life.

I appreciate all the input provided by Sandia National Laboratories researchers that I had the opportunity to work with. I especially would like to thank Dr. Matthew J. Reno for providing research direction and insightful ideas.

I am sincerely grateful to my housemates Miguel and Carlos, who have become my family away from home. Their positiveness, humor, hard work, and perseverance inspired me to keep pushing forward towards completing my thesis.

My deepest gratitude belongs to my parents Manuel and Puri for their love, trust, and encouragement during my entire life as a student, without whose unconditional support I would not be here.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

PV	Photovoltaics
VRE	Voltage Regulation Equipment
MW	Megawatt
kW	Kilowatt
QSTS	Quasi-Static Time Series
NAPS	North American Power Symposium
NREL	National Renewable Energy Laboratory
COM	Component Object Model
PMU	Phasor Measurement Unit
ACES	Advanced Computational Electricity Systems
OpenDSS	The Open Distribution System Simulator

SUMMARY

High penetration of solar photovoltaics can have a significant impact on the power flows and voltages in distribution systems. In order to support distribution grid planning, control and optimization, it is imperative for utilities to maintain a database of the location and sizes of PV systems. This research describes an efficient method to determine the location of solar PV systems in distribution circuits, based on voltage magnitude measurement streams.

The algorithm leverages the expected impact of solar injection variations on circuit voltages and takes into account the operation and impact of changes in voltage due to discrete Voltage Regulation Equipment (VRE). The estimation model enables recovering the most likely location of PV systems, as well as voltage regulator tap and switching capacitors state changes. The method has been tested on individual and multiple PV system, using the Chi-Square test as a metric to evaluate the goodness of fit. Simulations on the IEEE 13-bus and IEEE 123-bus distribution feeders demonstrate the ability of the method to provide consistent estimations of PV locations and VRE actions.

CHAPTER 1. INTRODUCTION

This chapter covers the motivation, objectives, and contributions to the research.

1.1 Motivation

Solar photovoltaic (PV) systems continue to be deployed at a rapid pace in distribution systems and at customers premises. Typical installations include distributed small residential systems in a neighborhood or subdivision, and medium-size central solar either utility-owned or owned by medium and small commercial and industrial customers. Databases derived from the solar PV interconnection requests and permits provide information about the location and size of PV systems. However, those databases are difficult to maintain and can have errors regarding the location of the system, the phase to which it is connected, the PV array orientation, etc. [1]. Even if the databases were accurate, an installed PV system may differ from its specifications due to the effects of damage, shading, or soiling, and may not be providing the expected power injections. The connection of unpermitted PV has also been reported [2]. Moreover, power injections from PV systems are usually not measured directly by utilities. According to the utility survey conducted by EPRI [14], 63% of utilities do not record the PV tilt and azimuth and 74% do not have any metering on residential PV systems.

Poorly maintained databases of PV systems can cause issues for distribution planning and operation. For instance, requests for interconnections that are analyzed based on inaccurate

circuit models can result in significant errors in the approval process and even the erroneous operation of controls in real-time [3,4]. Therefore, utilities are interested in maintaining and validating their PV installation databases based on the actual performance of the circuit. While the location, size, and parameters of the PV systems installed, as well as the development of validated circuit models are all desirable, this paper focuses on determining the location of PV systems through the analysis of voltage magnitude measurements.

The injection of active power from solar PV (or the corresponding decrease in a customer net demand) increases the voltages in the distribution circuit. Solar irradiance, power injections from PV systems, and hence the circuit voltages can vary significantly and rapidly depending on changes in cloud cover in a given region. The variations of voltages magnitudes seen at a particular node, depend on the size and also on the location of the PV system in the circuit. The growing availability of data streams from smart meters and distribution circuit sensors make it possible to estimate of PV locations based on voltage measurements. In particular, the work in [4] illustrates a quasi-linear relationship between injections of active power and circuit voltage magnitudes.

The injection of reactive power also impacts circuit voltages, but their sensitivities are substantially different from active power sensitivities [5]. The magnitude of voltage sensitivities in terms of reactive power increases with low power factors loads, especially

in feeders with low R/X ratios. That is, the impact of reactive power is more significant at medium voltage feeders than at low voltage feeders.

This research describes a method that leverages emerging data streams and voltage sensitivities based on the circuit model for estimation of solar PV location in distribution networks.

1.2 Objectives

This objective of this research is to predict the location of PV based on voltage magnitude measurements in distribution systems. A physics-based (i.e. non-black-box) data-driven algorithm is presented. This allows the distribution grid models to dynamically adapt to changing conditions on the grid for widespread applicability to all distribution feeders with monitoring.

The research pursues the following objectives:

- To estimate the location of any combination of 3-phase and 1-phase PV systems in a small unbalanced 3-phase distribution system. The location is specified by a given "bus" (for 3-phase PV) or a "node phase" (for 1-phase PV) on the distribution circuit.
- To determine the estimation confidence regarding the location.
- To study the impact of capacitors banks in the estimations.
- To study the impact of voltage regulators in the estimations.

- To develop a framework that enables the method to estimate at time points with voltage increments caused by the action of switching capacitors and tap changes.
- To verify that the method works for two different test feeders (IEEE13 and IEEE123).

1.3 Contributions and Acknowledgements

The research developed and reported in this thesis is based upon work supported in part by Sandia National Laboratories sponsored project to the Georgia Institute of Technology under the Solar Energy Technologies Office Program.

The research was conducted in the Advanced Computational Electricity Systems (ACES) lab at Georgia Tech, and was divided into 2 parts. During the first part, the ACES team was formed by Georgia Tech Master students Ahmad Khan, Jordan Sihno Mbeleg and Cristian Gómez Peces, under the guidance of Dr. Santiago Grijalva. The second part of the project was carried out by Cristian Gómez Peces under the supervision of Dr. Santiago Grijalva.

The first part of the research consisted of the core idea implementation proposed by Dr. Santiago Grijalva. Below are the team members contributions:

- Ahmad Khan: Extraction of voltage data streams using FastQSTS tool and implementation of the estimation for multiple single-PV simulations.

- Jordan Sihno Mbeleg: Literature review of current methods to estimate PV location, and implementation of the goodness of fit.
- Cristian Gómez Peces: Extraction of voltage sensitivities, study of negative sensitivities, and the impact of time window used to evaluate measurement vector.

This first part of the research project concluded with a paper submission to the North American Power Symposium (NAPS). The second part of the project involved further analysis on the PV location algorithm and the incorporation of VRE into the model, in order to prevent misclassifications and enable the algorithm to detect tap changes and states of capacitors. Below is the list of tasks carried out by Cristian Gómez Peces:

- Study of the accuracy of the method: proposal to use principal component to determine the algorithm's precision.
- Analysis of the number of degrees of freedom and minimum eigenvalues of the gramian of the sensitivity matrix.
- Develop low-level framework to automatically extract voltage sensitivities and synthetic voltage data streams using GridPV toolbox.
- Extraction of voltages measurements given multiple PV case scenarios and study multiple PV location estimations.
- Development of methodologies to include the impact of VRE in the model estimation.
- Study of limitations of the algorithm.

- Analysis of the performance of the method given missing voltage data and noise in the measurements.
- Expand and test the method implementation to a larger distribution network (IEEE 123-bus test feeder).

Most of the tasks that were carried out in the first part are presented in Chapter 3, while the rest of topics is discussed throughout Chapters 2, 4 and 5.

CHAPTER 2. BACKGROUND AND ASSUMPTIONS

This chapter covers the literature review and state of the art in terms of PV location estimation. The assumptions required for the algorithm to work are discussed. The distribution feeder models are introduced, in addition to the tools used to obtain the voltage sensitivities and data stream measurements.

2.1 Previous Studies to Estimate PV Locations and VRE actions

Several approaches have been proposed in the literature to determine the PV generation based on data-driven methods. The research in [2] uses net meter data to detect PV installations and their size. The study reports a size estimation error of 4.176% for a 5kW PV system when local cloud coverage information is included in the estimation. The work in [6] presents a deep neural network approach for estimating PV size, tilt, and azimuth using behind-the-meter data. The method estimates PV size with an error of 2.09% in a data set with fixed tilt and azimuth values and 3.98% in a data set with varying tilt and azimuths, and it is reasonably robust to erroneous training data.

The work in [7] proposes a clustering-based method, which does not require labelled training information or exogenous data, and results in high classification accuracy. References [8] and [9] describe machine learning algorithms to detect solar PV arrays using high resolution color satellite imagery. Support vector classification, support vector machine and bootstrap methods are used in [10] to detect whether a customer has a distributed PV system and to predict its output characteristics. A method to nowcast power generation of behind-the-meter PV systems, by modeling a correction between harmonic

current injection and PV injection using artificial neural networks is described in [11], and a method that uses the correlation between the net load of a bus line and solar radiation intensity is described in [12]. The overall literature however lacks a method that can estimate the location of PV based on voltage measurements.

The study in [14] uses an approach to exploit the high correlation of diurnal and nocturnal demands by using Gaussian mixture models and maximum likelihood estimation-based techniques to disaggregate customer-level behind-the-meter (BTM) PV generation. An unsupervised framework is presented in [15] for joint disaggregation of the net load readings into the solar PV generation and electric load, where estimations are made based on a mixed hidden Markov model (MHMM).

The work in [1] presents a study of the PV injection, tap changes and switching capacitor impact on the system voltages, and [18] also defines the impact in terms of sensitivity planes. These core-fundamental concepts are leveraged in this research to estimate active power injections, and VRE actions.

Various methods have been proposed to estimate the tap setting of voltage regulators. In [19], the authors introduce a slight modification into the power system state estimation to predict both the voltage turns ratio and the phase-shift angle of transformers. The authors in [20] present a method to estimate the tap positions based on the residuals from the state estimation, incorporating prior information about measurements and tap positions to increase robustness when dealing with bad data. Despite this, there are no methods that use sensitivity analysis to estimate the tap changes given a voltage measurement stream. This

paper presents a novel combination of the PV and VRE sensitivity analysis model to detect impacts on the circuit voltages.

2.2 Study Assumptions

The method works under certain conditions. Since distribution feeders are non-linear systems, the sensitivities are obtained by linearizing the dynamic system at a certain point of operation. The sensitivities of voltage magnitude with small injections of PV power are fairly linear if the voltage regulating devices remain in the same state. According to [3], the voltage shall be in the range from 0.95 to 1.05 p.u. in order to ensure an optimal operation of the method.

The proposed method can estimate the location of PV systems not solely based on the voltage magnitude measurements, but also requires an accurate distribution system model and voltage sensitivities. In other words, the method assumes a perfect model of the distribution feeder. In reality, distribution models have errors and inaccuracies that may impact the estimation. The impact of these inaccuracies is studied in Chapter 5.

The proposed model to estimate PV injections only supports discrete voltage regulating devices. Consequently, the impacts of continuous voltage regulating devices such as advanced inverters (e.g. with volt-var regulation) are not taken into account in this study.

The method works under the hypothesis that a PV system will leave a unique footprint in terms of voltage increase depending upon the node to which it is connected. Depending on

the distribution circuit, some nodes may have very close voltage sensitivities. These similarities can cause singularities in the estimation framework, so it is assumed that the PV location candidates have different sensitivities. This is also the reason why one of the nodes connected by a switch is eliminated from the sensitivity matrix, since witches are defined with a very low impedance. Thus, both locations will have the same voltage sensitivity when the switch is closed, and the PV system is placed in either of both locations.

As the solution proposed takes the form of a state estimation, we require more states than measurements in order to get accurate results, i.e. we assume that number of PV location candidates will be lower than the number of measurements available. The impact of the number of degrees of freedom is discussed in section 3.4.2, and the impact of missing voltage data will be introduced in section 5.3.

2.3 Distribution Feeder Models

The research has been conducted using two test feeders. The simulations on the IEEE 13-bus radial distribution feeder are useful to gain insight, debug and troubleshoot the first implementations of the algorithm. In a later the stage of the research, the implementation of the method was expanded to a larder test feeder: the IEEE 123-bus distribution feeder to show the method's functionality on a larger and more realistic test feeder. In this section, the characteristics of each circuit are presented.

2.3.1 The IEEE 13-bus feeder

This circuit operates at 4.16 kV, has unbalanced loading, is relatively short and is highly loaded. This circuit also has a single voltage regulator at the substation, overhead and underground lines, shunt capacitors, and one in-line transformer. The circuit topology is presented in Figure 1.

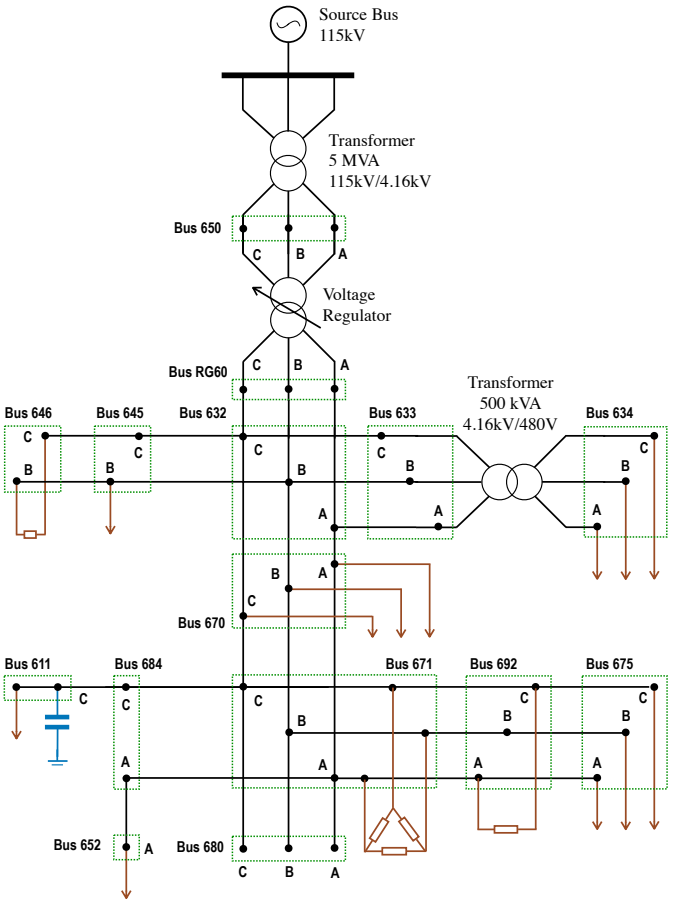


Figure 1 – Diagram of the IEEE 13-bus system.

This distribution system contains buses that have 1, 2 and 3 phases. It has a total of 32 nodes (bus-phase combinations). It is assumed that meters that can measure the voltage magnitude are located at each one of the 32 circuit nodes. Buses 671 and 692 are connected through a closed switch and hence their phases have the same voltage magnitude. The loads are provided with the same load time profile.

2.3.2 The IEEE 123-bus feeder

The IEEE 123-bus distribution system operates at 4.16 kV and has an unbalanced loading. The circuit has a several voltage control devices in addition to the regulator located at the substation: a 1-phase regulator between nodes 9 and 14, 2 1-phase regulators between 25 and 26, and 3 1-phase voltage regulators between nodes 160 and 67. The system presents switches that have been closed to keep a radial topology. The circuit topology is presented in Figure 2.

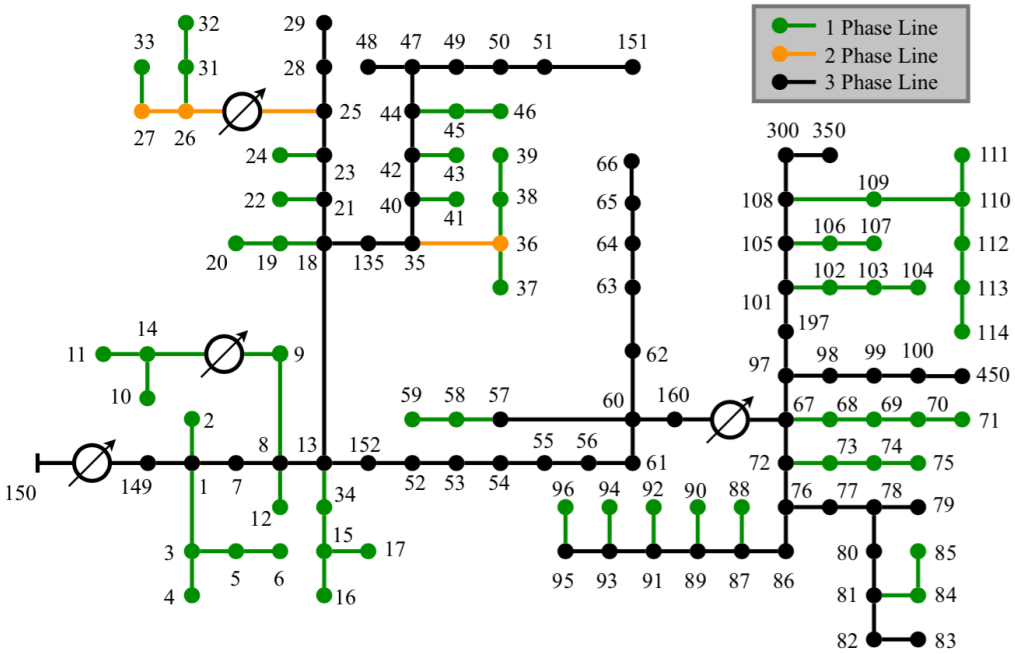


Figure 2 – Diagram of the IEEE 123-bus system.

2.4 Synthetic Data Generation and Tools

The Open Distribution System Simulator (OpenDSS) is used to run power flows and quasi-static simulations required for this research study. It is an electrical system simulation tool for electric utility distribution systems. The program supports nearly all RMS steady-state (i.e., frequency domain) analyses commonly performed for utility distribution systems planning and analysis. In addition, it supports many new types of analyses that are designed to meet future needs, many of which are being dictated by the deregulation of utilities worldwide and the advent of the Smart Grid. Many of the features found in the program were originally intended to support distributed generation analysis needs. Other features

support energy efficiency analysis of power delivery, smart grid applications, and harmonics analysis [23].

What makes OpenDSS especially useful for this research is its integration with MATLAB through the COM object. From MATLAB, it is possible to initialize the object and point it to OpenDSS to set up simulations and retrieve specific data from the power flows. In addition, the ACES lab developed several tools to manage the COM object in a high-level, and control and retrieve data in a straightforward way. In particular, this research heavily uses the GridPV toolbox, developed by Matthew J. Reno and Kyle Coogan [25].

The circuit models introduced in section 2.3 are available for retrieval at SourceForge, the official business software platform for OpenDSS. The models have been modified to accommodate the needs of the research. This includes reducing the operating band of voltage regulators to see how more frequent VRE actions, or modifying the load of the system to make more balanced and remain within voltage range assumption mentioned in section 2.2.

The PV profile is based on irradiance data provided by NREL and represents the actual irradiation values observed on January 1, 2011 in Oahu, Hawaii [21]. This file is defined in a SNG format to be run in duty mode by OpenDSS. The PV profile selected has a 10-second resolution for one day, which represents a total of 8,640 data points. It is observed that highest values of PV irradiance as well as the highest changes in the PV profile

occurred between 11am and 1pm. Figure 3 includes a typical PV profile during a intermittent cloudy day.

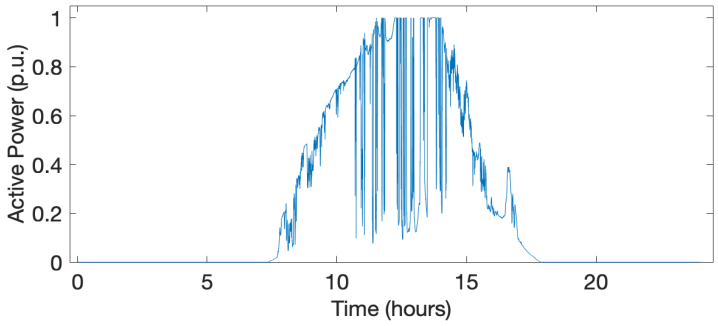


Figure 3 – PV profile used in simulations.

2.4.1 Automatic Voltage Measurements and Sensitivities Generation Tool

To run multiple experiments and test whether it is possible or not to predict the PV location at different nodes, it is necessary have not only all voltage sensitivities but also synthetic voltage measurements data to get the measurement vectors and make estimations. During this research, a tool that automatically obtains the sensitivities in the system for each of the accessible nodes, capacitor banks and voltage regulators was developed. In addition, it also gathers voltage data stream measurements that are obtained from placing different PV case scenarios across the feeder. This tool has several advantages. First, it is approximately 15 times faster than FastQSTS, the previous tool that was being used. FastQSTS provides so much data that is not relevant for this application, such as power flows, which made it

slower. Second, it let us configure case scenarios with no PV or voltage control activated, which is required to calculate the PV sensitivities. Third, it is straightforward to specify the PV scenarios we want to run via Excel file. Switching capacitor banks and voltage regulators can be enabled via excel files as well. These files include the nodes in which the PV and switching capacitors will be placed at each simulation, and the configuration of the voltage control devices. A program in MATLAB was written to take all this input information and communicate with OpenDSS to run the quasi-static time-series simulations. Figure 4 shows the workflow of the application.

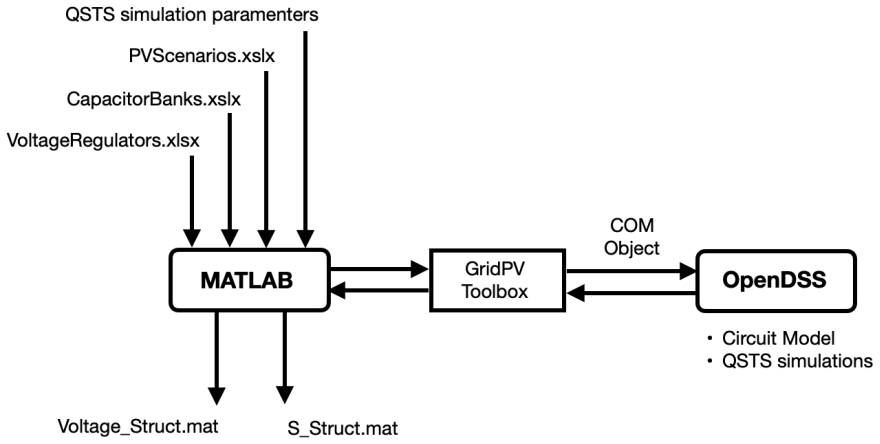


Figure 4 – Workflow of Voltage and Sensitivities Generation Tool

The program will run 2 sets of simulations. First, it will obtain the voltage profiles running QSTS simulations with the control devices activated. The voltage recorded in each simulation will be saved in a structure called Voltage. The sensitivities are obtained as follows. First, a power flow is run with no PV and VRE disabled. This simulation is called

the base case scenario. For each location selected on the PVscenarios.xls file, a 100kW PV is placed supplying its maximum power. The sensitivity vector will be exactly the increment of voltage that is obtained when we compare to the base case scenario. The same procedure is applied to get capacitor and tap changes sensitivities: we calculate the voltage increase that comes specifically from a state change or tap position change, respectively.

2.4.2 Sensitivities Visualization Tool

A visualization tool was created to visualize the sensitivities on the IEEE 13-bus feeder diagram. This program takes as input the sensitivity vector and displays it over the circuit scheme in terms of bubbles. Figures below show some examples of PV sensitivities.

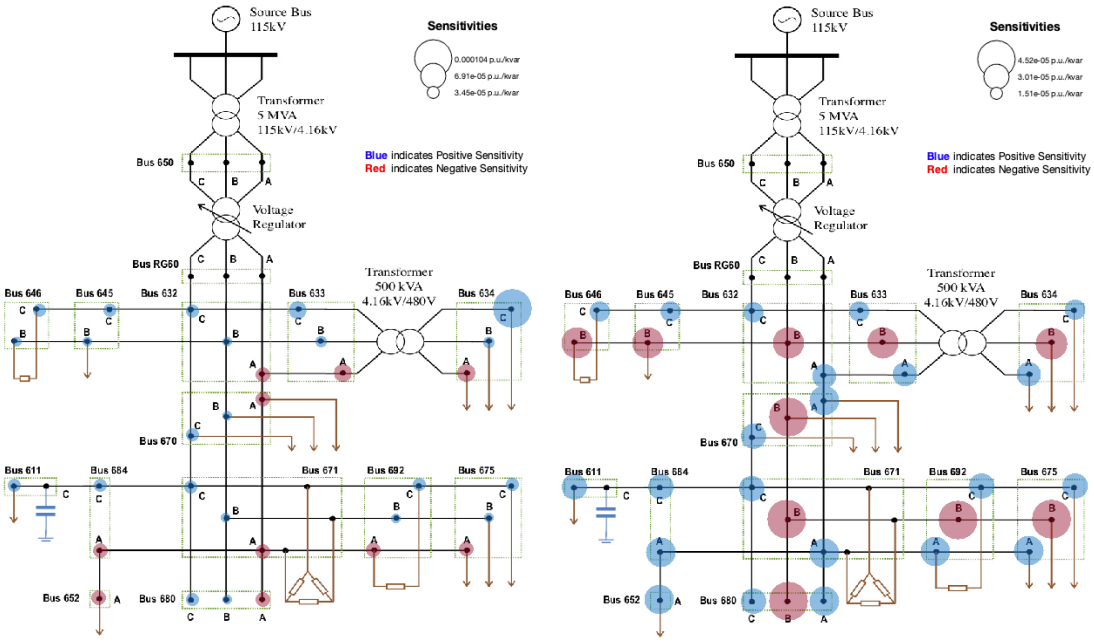


Figure 5 – Voltage sensitivities for PV at nodes 634C (left) and 670A (right).

CHAPTER 3. VOLTAGE SENSITIVITY METHOD

This proposed voltage sensitivity method is based on 3 key elements: calculation of the voltage sensitivity matrix, estimation method using measured voltage, and goodness of fit. The following sections explain how to compute these elements, what impacts the accuracy of the method, and shows estimation examples with multiple PV case scenarios.

3.1 Calculation of the Voltage Sensitivity Matrix

In [3,4], it was established that the injection of PV active power at a given location results in changes in the voltage magnitudes and that these changes are fairly linear and consistent along time periods with solar irradiance variations. Let *i* be the node (the electric point) corresponding to a phase *p* of bus *b.* The set $\mathrm{N} = \{1, \dots, i, \dots N\}$ contains all such nodes in the system. We denote the change in voltage magnitude of node *i* with respect to a power injection change at a PV location ℓ as:

$$s_{i\ell} = \partial V_i / \partial P_\ell \tag{3.1}$$

The *N* node voltages in the system change due to this power injection. Thus, a vector of sensitivities with respect to injection at location ℓ can be written as $\mathbf{s}_\ell = \partial \mathbf{V} / \partial P_\ell$, where **V** is the vector containing all the node voltage magnitudes in the circuit. One can obtain the values of the vector $\mathbf{s}_\ell$ for a given location ℓ by the following steps:

1. Solve the three-phase unbalanced power flow of the circuit for a baseline condition without the PV system,
2. Solve a second power flow with the PV system installed at location l and all voltage regulating equipment (VRE) disabled, and
3. Record the voltage magnitude differences at each node by comparing the power flows with and without PV.

If the PV systems considered are known to be 3-phase, then the possible locations correspond to all the 3-phase buses in the circuit. If the PV systems considered are 1-phase, then the space of possible locations corresponds to all the nodes in the circuit. Let us denote by $\mathcal{L} = \{1, \dots, \ell, \dots L\}$ the set of all possible PV locations. By solving a power flow for the baseline and one power flow with the PV at each location ℓ, one can determine an $N \times L$ sensitivity matrix of node voltage changes with respect to PV injections at the L locations:

$$\boldsymbol{S} = \partial \mathbf{V} / \partial \boldsymbol{P} \tag{3.2}$$

We note that the voltages at the slack node do not change, and that the changes in voltages at nodes connected by switching devices are identical. The matrix S is full rank if the matrix column corresponding to nodes in the slack bus and one of each pair of nodes that are terminals of switches have been eliminated.

3.2 Estimation Method Using Measured Voltage

Smart meters and other sensors being deployed in the grid can measure voltage magnitude over time for each phase. Most smart meters deployments record measurements every 15 or 60 minutes. In this paper, we assume that the utility is recording voltage magnitude at various nodes in the system, that this data stream has a fixed granularity *G*, and that data records are available for some time horizon *H*. The resulting measurement data streams has size $M' = H/G$.

Figure 6 illustrates the node voltage data stream for one day for a small system with one medium-size solar PV system. The voltages change significantly during the day due to the presence of the PV system. One can also observe that during this day the PV power injection has experienced significant variability.

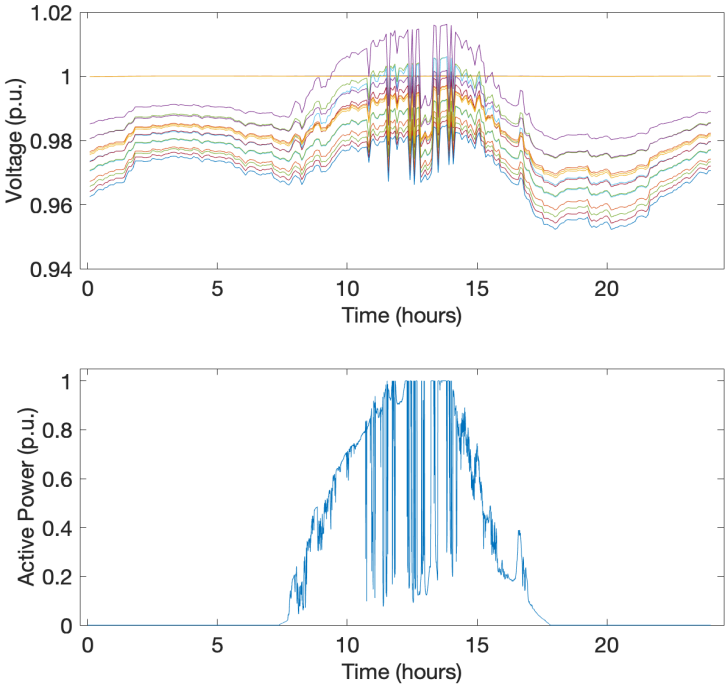


Figure 6 – Voltage and active power injected as a function of time in the IEEE 13 bus system. Plots illustrate the impact of demand and solar PV.

Let us consider an $N \times M'$ matrix Z that contains the data stream of all the measured node voltage magnitudes for a given time horizon H. We are interested in the change of voltage magnitudes over time (as a function of the changes of PV injections). By taking the simple difference from one measurement scan to the next, we can obtain a matrix of voltage differences $\boldsymbol{D}'$ of size $N \times (M' - 1)$. During, the night the changes in voltage due to solar PV are zero. Thus, we select intervals of measurements during the day, where solar PV power magnitude as well as the variation in power is likely to be significant. This subset of voltage differences is denoted by a matrix $\boldsymbol{D}$ of size $N \times M$. In the trivial case of a single point, the matrix $\boldsymbol{D}$ corresponds to a vector d of size $N \times 1$. The vector d contains the

measured changes in voltage magnitude due to the PV injection change at a given location ℓ. Thus, the vector $\boldsymbol{d}$, down-scaled by the size of the PV system must be equal to one column of matrix $\boldsymbol{S}$, the exact column of sensitivities corresponding to that location ℓ. In other words, it must be true that if the PV is installed at location ℓ, then:

$$\boldsymbol{s}_\ell = \partial \mathbf{V} / \partial P_\ell = \frac{1}{\alpha} \boldsymbol{d} \tag{3.3}$$

To further illustrate the meaning of α, let us assume that matrix $\mathbf{S}$ was obtained by sequentially simulating a 100kW PV at each location ℓ. Then the units of the sensitivities $s_{i\ell}$ are in volts per 100kW. Since the vector d has a unit of volts, the units of α are 100s of kilowatts, i.e. $\boldsymbol{s}_\ell = \boldsymbol{d}$. Then, the change in PV power injection that resulted on changes in voltages equal to $\boldsymbol{d}$ was 100kW. Note that this means that the size of the PV system must be at least 100kW, e.g. it is not possible to determine the size of the PV system since a variation from zero to full PV capacity is unlikely to occur from one scan to the next.

The measurements obtained from actual sensors will unavoidably contain errors due to sensor class, model inaccuracies, etc. Thus, the scaled vector $\boldsymbol{d}$ will be close to $\boldsymbol{s}_\ell$, but not exactly equal. Each column of matrix $\boldsymbol{S}$ represents the "direction" of the changes in voltage. The location of the PV system can then be determined by finding the column that is best aligned with the direction of the measured vector $\boldsymbol{d}$. We want to estimate a vector $\mathbf{x}$ such that $\boldsymbol{Sx} = \boldsymbol{d}$. This problem is known to have a unique least-squares solution:

$$\hat{\boldsymbol{x}} = (\boldsymbol{S}^{\boldsymbol{T}}\boldsymbol{S})^{-1}\boldsymbol{S}^{\boldsymbol{T}}\boldsymbol{d} \tag{3.4}$$

Since the columns of $\mathbf{S}$ are linearly independent, $(\mathbf{S}^{\mathbf{T}}\mathbf{S})^{-1}$ is computable. The vector $\hat{\boldsymbol{x}}$ is the projection of d onto the subspace $\boldsymbol{S}$. It is the minimization of the components of $\boldsymbol{d} - \boldsymbol{S}\boldsymbol{x}$, such that $\|\boldsymbol{d} - \boldsymbol{S}\hat{\boldsymbol{x}}\| \leq \|\boldsymbol{d} - \boldsymbol{S}\boldsymbol{x}\|, \forall \boldsymbol{x} \in \mathbb{R}^L$.

In order to incorporate not only one point, as in vector d, but more information available from the data stream, one can obtain a metric that captures the changes in voltage magnitude for a given period during the day. A suitable metric is the sum of the M_{pos} positive values of changes in voltage from matrix **D**. Thus instead of vector d in (4), we use the following vector:

$$\bar{\boldsymbol{d}} = \frac{1}{M_{pos}} \sum_{t=1}^{T} pos(\boldsymbol{d}_t) \tag{3.5}$$

3.3 Goodness of Fit

Let us assume again that we have a single vector of voltage magnitude deviations $\boldsymbol{d}$ obtained from the difference of voltage measurements at two points in time, resulting on an estimated value $\hat{\boldsymbol{x}}$. The vector of estimated voltage differences at each node is given by:

$$\hat{\boldsymbol{d}} = \boldsymbol{S}\hat{\boldsymbol{x}} \tag{3.6}$$

The normalized residuals of the voltage differences d_i are assumed to have a normal distribution $r_i \sim N(0,1)$, where: $r_i = (\hat{d}_i - d_i)/\sigma_i$. Voltage meters and smart meters

usually have an error of less than 1%. In this paper, we assume that $\sigma_i = 0.01$.

The least squares solution $\hat{\mathbf{x}}$ minimizes the sum of the squares of r_i:

$$\sum_{i=1}^{M} s_i^2(\boldsymbol{x}) = \chi^2 \geq \zeta = \sum_{i=1}^{M} s_i^2(\hat{\boldsymbol{x}}) \tag{3.7}$$

We note that the value of ζ can alternatively be computed as:

$$\zeta(\hat{\boldsymbol{x}}) = [\boldsymbol{S}\hat{\boldsymbol{x}} - \boldsymbol{d}]^T \boldsymbol{\Omega}^{-1} [\boldsymbol{S}\hat{\boldsymbol{x}} - \boldsymbol{d}] \tag{3.8}$$

where $\boldsymbol{\Omega}^{-1}$ is a diagonal matrix with entries $1/\sigma_i$. The probability that the above event $\chi^2 \geq \zeta$, is given by the chi-square distribution:

$$\Pr[\chi^2 \geq \zeta] = 1.0 - \Pr[\zeta, \nu] \tag{3.9}$$

where $\nu = M - L$ is the number of degrees of freedom. $P = \Pr[\chi^2 \geq \zeta]$ represents the confidence level of the PV injection being at the estimated location. The smaller the value of ζ, the better the estimation will be. If a data stream is used instead of a single difference measurement, then the vector $\bar{\mathbf{d}}$ should be used in the above equations.

In summary, the goodness of fit is determined by the following process:

1. Compute the estimate $\hat{\boldsymbol{x}}$ using $\bar{\boldsymbol{d}}$ in equation (3.4)

2. Compute the objective function ζ using equation (3.8)

3. Compute the probability $\Pr[\chi^2 \geq \zeta]$ using equation (3.9)

3.4 Accuracy Study

The IEEE 13-bus test feeder is used to analyze the method accuracy. We start with the estimation of 3-phase PV systems, which inject the same amount of power at each phase. The possible locations of the PV system correspond to the 3-phase buses: 633, 671, 670, 675, and 680 in Figure 1.

The estimation process starts with determining the matrix **S** by connecting PV systems sequentially at each 3-phase bus and recording the changes in voltage magnitude in the 32 nodes in simulation. We assume that the injection of PV power has a power factor of 1.0. This matrix is of size 5 locations times 32 nodes. Figure 7 (top) illustrates the columns of the sensitivity matrix **S**. Each one of the bar series for the five locations represents a unique signature on how the power injection at that locations changes the voltages in each one of the 32 circuit nodes. In order to be able to obtain solution for the estimates, **S** must be full rank.

We conducted the test by using only the voltage measurements in the time range from 11 am to 1 pm, since this is the time of the day when solar PV output is usually the highest. We assume a 5-minute (300 seconds) resolution of voltage measurements. In order to form vector $\bar{\boldsymbol{d}}$, we selected sub-intervals of 10 minutes. With this vector $\bar{\boldsymbol{d}}$, we computed the

estimate of vector $\hat{\boldsymbol{x}}$ for each interval. The voltage variation caused by solar PV variation during an interval allows us to pinpoint the location of the PV.

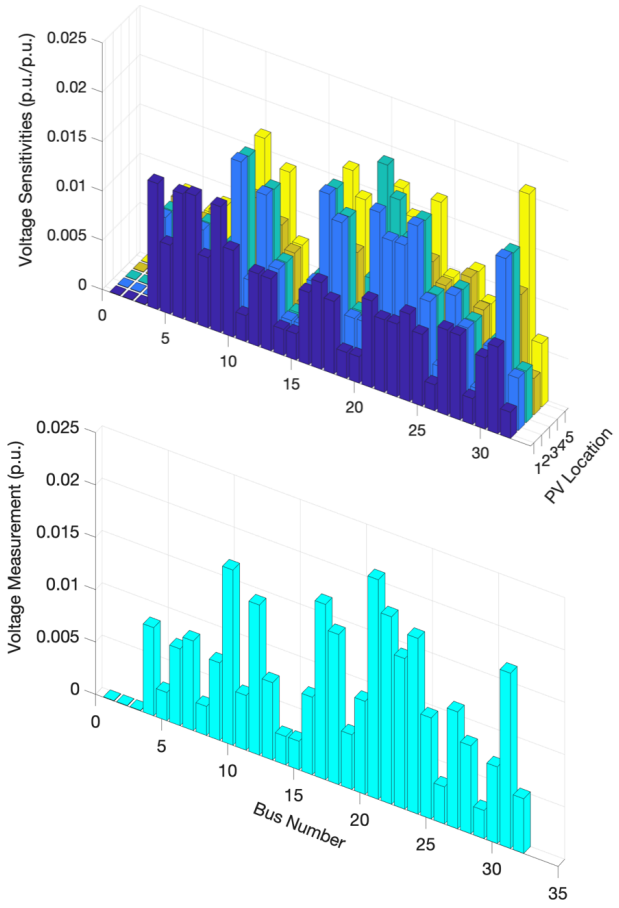


Figure 7 – Representation of Sensitivity Matrix (top) and measurement vector (bottom).

As an example of the expected relation between the voltage sensitivities and measurements vector, let us consider the voltage measurements in the range from 12:50pm to 1:00 pm.

Figure 7 (top) represents the expected changes in voltage magnitude (the matrix **S**), while Figure 7 (bottom) presents the actual vector $\overline{\boldsymbol{d}}$ obtained using equations (3.4) and (3.5). We observe that the shape of the vector $\bar{\mathbf{d}}$ is closely resembles the shape of the column of matrix **S** corresponding to location number 3. Thus, the PV must be located at location 3. The size of the PV system modelled to generate the matrix $\boldsymbol{S}$ is 1000 kW. Table I presents the values of the estimated vector $\hat{\boldsymbol{x}}$ for the estimated locations versus actual locations during the time period from 11:30 to 11:40am. The high values in the diagonal indicate that the estimation is correct. Specifically, a high value close to 1.0 means that the PV system is highly likely to be located at that bus, while a value closer to zero means that the PV is highly unlikely to be located at that bus.

Table 1 – Values of $\hat{\mathbf{x}}$ for Estimated and Actual Locations from 11:30am to 11:40am

Estimated Location	Actual Location				
	633	671	675	670	680
633	1.138	0.0205	0.0205	0.0205	0.0205
671	0.0401	1.158	0.0401	0.0401	0.0401
675	0.0173	0.0173	1.135	0.0173	0.0173
670	0.0069	0.0069	0.0069	1.125	0.0069
680	-0.0316	-0.0316	-0.0316	-0.0316	1.086

Another way to graphically illustrate the estimation is through the use of a spider plot as shown in Figure 8. The outer labels of the spider plot represent the actual PV locations, while the colored subplots represent the estimated locations. The spider plot is a graphical representation of the vectors $\hat{\boldsymbol{x}}$ for each simulation. The more a subplot in the spider plot tilts towards an outer label, the higher the probability of a PV system at that location. The more the plot bends closer to the inner zero points, the less likely it is that the PV is present at that location. As can be seen, Figure 8 consists of near perfect shapes. Each subplot has almost zero value for non-PV bus locations, while they have a significantly higher value for the PV bus. The estimation is highly accurate and consistent for the five locations, resulting in a clear pentagon being formed by joining the points associated to the highest estimated values for vector $\hat{\boldsymbol{x}}$.

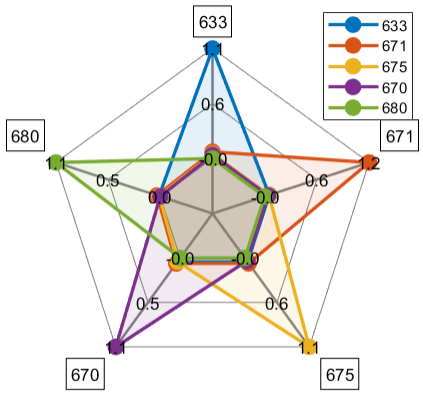


Figure 8 – PV Location Prediction from 11.30am to 11.40am. Graphical representation of values of x̂ from Table 1.

The goodness of fit for the results of both estimations are presented on Table II, according to equations (3.8) and (3.9).

Table 2 – Statistical Results for 3-phase PV Estimations

PV Location	633	671	675	670	680
	Simulation 1: from 11.30am to 11.40am				
ζ	0.0517	0.0517	0.0517	0.0517	0.0517
$\Pr[\chi^2 \geq \zeta]$	0.9997	0.9997	0.9997	0.999	0.999
	Simulation 1: from 12.20pm to 12.30pm				
ζ	0.0047	0.0047	0.0047	0.0047	0.0047
$\Pr[\chi^2 \geq \zeta]$	1.0	1.0	1.0	1.0	1.0

We repeat similar estimations for thirty different ranges between 11am and 1 pm. We obtain the values of the estimated voltage differences $\widehat{\boldsymbol{d}} = \boldsymbol{S}\widehat{\boldsymbol{x}}$, using equation (3.6). We then compute the objective function of the estimation, i.e. the sum of the normalized residuals. Figure 9 presents the value of the objective function for the thirty estimations developed in this manner. This value of the objective function $\zeta(\widehat{\boldsymbol{x}})$ is compared with the value of the χ^2 function at 95% confidence for M-N degrees of freedom, which is equal to 40.11. In all the estimations, the objective function is significantly smaller, indicating a good fit. This shows that the algorithm works seamlessly over a range of different time points in the day. The lowest values of the objective functions occur when estimations are performed for intervals that display the highest solar PV magnitude and variation, since this allows the matrix **S** to better represent voltage changes due to PV variations. For time

intervals of low PV magnitude and variation, the voltage sensitivities consist mostly of voltage changes due to changes in system demand. After analysis of the irradiance data used for this study, it was found that the highest values of PV magnitude and variation occur at different time intervals between 11am and 1pm. It will be important to locate the range of time during which solar PV produces higher voltage variations, since the algorithm will be more effective during these time frames.

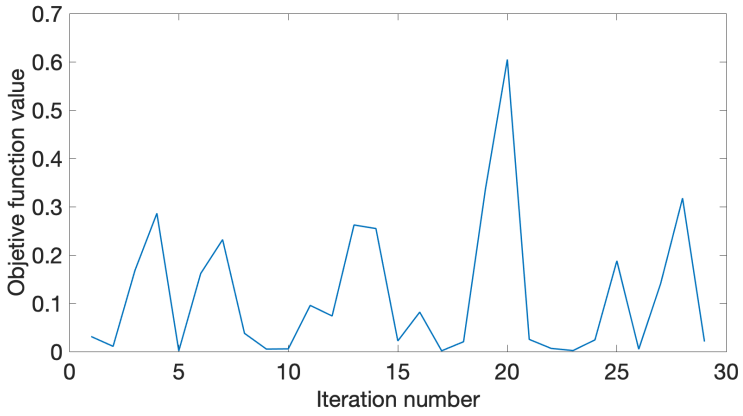


Figure 9 – Objective Function for 30 trials in different intervals.

For the simulation of 1-phase PV systems, we assume a granularity of 150s, and measurement vectors are computed using 5-minute time intervals. We used a PV size of 2000 kW for these simulations. These new values have been selected to test the method's robustness for varying values of timestep and PV size.

There are 32 1-phase nodes in the system. As with the 3-phase analysis, we avoid the slack bus, one of the two buses at the ends of the switch, and the regulated buses to have a total

of 20 possible PV locations to be analyzed. For the 1-phase analysis, we perform the same analysis for each PV location. We conduct the test for each phase separately for phases A, B and C to reduce error in the calculations. It is assumed that the phase information available to the utility is reasonably accurate.

Figure 10 shows the results of the analysis for PV systems on phase A of different buses. The results accurately predict the location of PV for each simulation. For phase A, buses 32, 633, 652, 670, 671, 675 and 680 are the possible PV locations. Similar simulations are performed for all phase B and phase C PV locations in the system. Figures 9a and 9b show the result of these analyses. For phase B, buses 633, 646, 670, 675 and 680 are possible locations, while for phase C, buses 611, 633, 645, 646, 670, 671, 675 and 680 are possible PV locations.

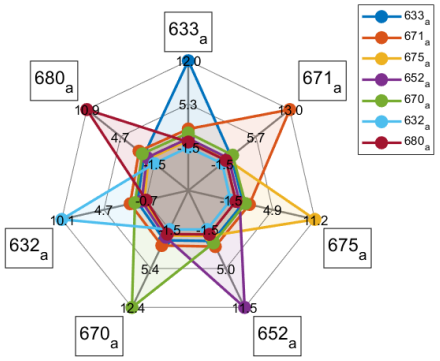
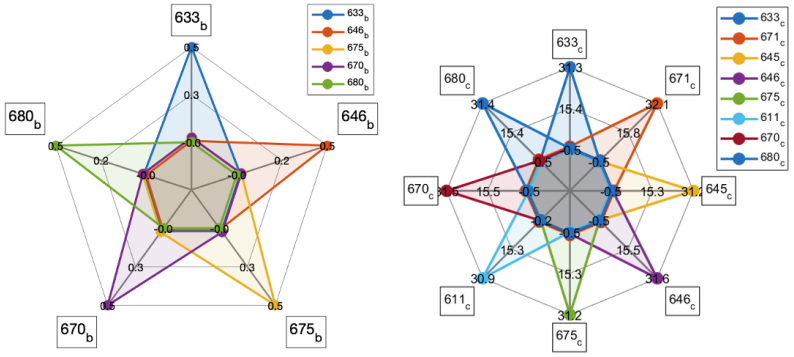


Figure 10 – PV Location Prediction (Phase A) at 12.20pm



a) Phase B at 12.45p. **b) Phase C at 11.25am**

Figure 11 – PV Location Prediction for phase B and C.

The figures show that the method accurately predicts the location of 1-phase PV using voltage sensitivities. The goodness of fit values for the 1-phase PV location estimation are presented in Table III, which shows the strong performance of the method.

Table 3 – Statistical Results for 1-phase PV Estimations

Phase	A		B		C	
Location	ζ	$\Pr[\chi^2 \geq \zeta]$	ζ	$\Pr[\chi^2 \geq \zeta]$	ζ	$\Pr[\chi^2 \geq \zeta]$
611					1E-05	1.00
632	3E-04	1.00				
632			5E-04	1.00		
633	3E-04	1.00	5E-04	1.00	1E-05	1.00
645					1E-05	1.00
646			5E-04	1.00	1E-05	1.00
652	3E-04	1.00				
670	3E-04	1.00	5E-04	1.00	1E-05	1.00
671	3E-04	1.00			1E-05	1.00
675	3E-04	1.00	5E-04	1.00	1E-05	1.00
680	3E-04	1.00	5E-04	1.00	1E-05	1.00

3.4.1 Metrics

The confidence level provides a metric that indicates whether the PV injection estimation is the cause of the measured increment of voltage or not. However, it does not provide any insight on whether the estimation was accurate or not. Two additional metrics are

considered to evaluate the classification: accuracy and precision. The accuracy is defined as the percentage of total success in multiple classifications:

$$Accuracy(\%) = \frac{True\ Positives + True\ Negatives}{Total\ Estimations} \tag{3.10}$$

The precision provides information about how clear the estimation is to identify the most likely PV location. In the case of a single PV system in the feeder, the ideal result would be 0 kW estimated for all nodes except for the node that is injecting power with the PV system. We use the principal components v_i of the estimation vector to evaluate the precision:

$$v_i = \frac{x_i}{\sum_{i\in\Omega} x_i} \qquad \forall\, x_i \geq 0 \quad and \quad v_i = 0\ if\ \ x_i < 0 \tag{3.11}$$

Figure 12 shows the graphical representation of the estimation. Each component of the estimation vector is actually a coordinate of the measurement vector projection onto the span of the sensitivity hyperplane, such that $\widehat{\boldsymbol{d}} = \boldsymbol{s_1} \cdot x_1 + \boldsymbol{s_2} \cdot x_2$ and $\boldsymbol{d_{error}}$ is minimized. The ideal case occurs when d belongs to the sensitivity plane and the confidence level is hence 100%, that is, the increment of voltage is completely due to PV systems. In the case of a single PV, the projection of the measurement vector should be alienated with one of the sensitivity vectors. In that case, the principal component v_i associated to that vector becomes non-zero, while the rest should yield to a considerably lower value.

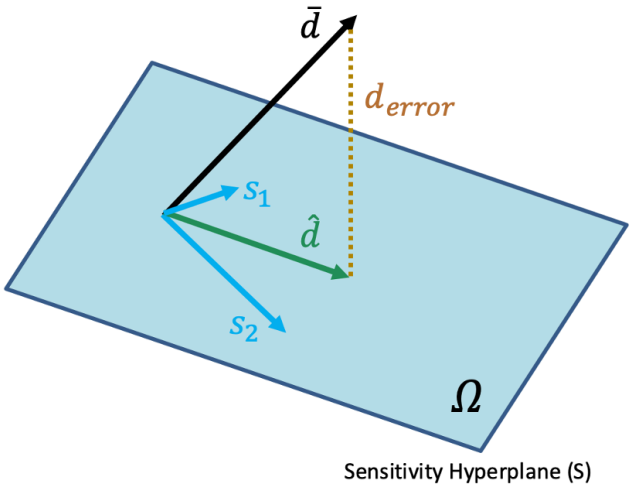


Figure 12 – Graphical representation of the state estimation method.

The proportion of principal components of vector **x** provides a metric bounded between 0 and 1. The components associated to higher proportions of explained variance are selected until up to an arbitrary threshold. This threshold can be set manually as percentage or use the scree plot to find the elbow. Usually, an explained variance of 70% is accepted by the overall literature. If we have a single PV in the feeder and consider an ideal case, we should get 100% for the PV estimated component. If we have 2 PV, we should get 50% at each component. With 3 PV, 33% at each component, etc. In fact, the metric is providing the relative PV injection across all PV location candidates. That is, if there are 2 PV in the feeder and one has double capacity, the resulting precision shall be close to 33% and 66%, respectively. Figure 13 shows a set of estimations for multiple simulations that included a

single PV. As can be observed, this approach provides a useful metric to evaluate the precision of the algorithm.

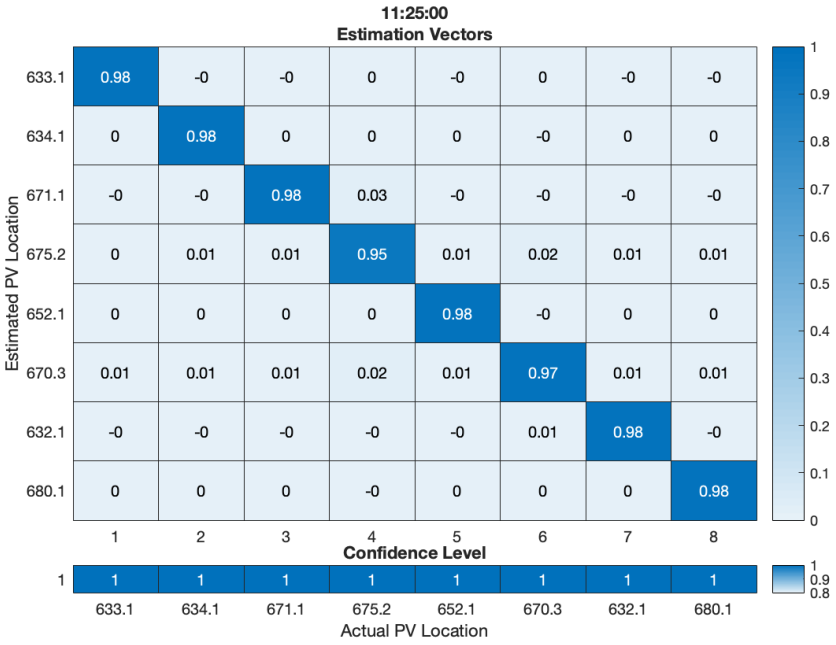


Figure 13 – Example of estimation using proportion of principal components.

The following sections analyze the impact of degrees of freedom in matrix S and the time frame for the measurement vector given the metrics defined above.

3.4.2 Impact of degrees of freedom on the estimation

The matrix $\boldsymbol{S}$ may introduce inaccuracies in the estimation due to the similarities of the sensitivity vectors. In particular, these inaccuracies are introduced when the inverse of the gramian of S is computed. $\boldsymbol{S}^T\boldsymbol{S}$ is called the Gramiam matrix of the columns of S. That is, it is the matrix of mutual scalar products of the columns of $\boldsymbol{S}$. Thus, $\boldsymbol{S}^T\boldsymbol{S}$ carries

information about pairwise relations of the columns of $\boldsymbol{S}$. For instance, the columns are mutually orthogonal if and only if $\boldsymbol{S}^T\boldsymbol{S}$ is diagonal, and they are linearly dependent if and only if $\boldsymbol{S}^T\boldsymbol{S}$ is singular. As a consequence, adding states will lead to lower minimum eigenvalues and the cause of singularities. The extreme case occurs when we add vector fully linearly dependent from $\boldsymbol{S}$. Figure 14 shows the value of the minimum eigenvalue as we add sensitivity vector into the sensitivity matrix.

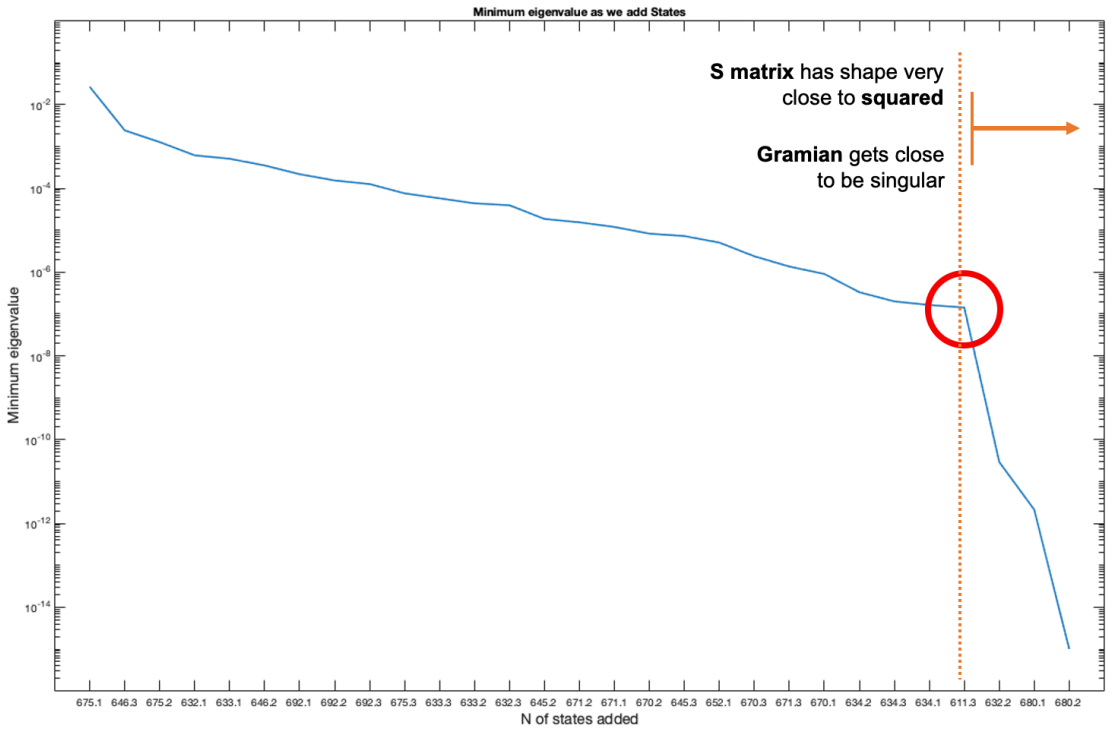


Figure 14 – Minimum eigen values of sensitivity gramian

As the sensitivity matrix adopts a squarer shape, the minimum eigenvalue decreases. However, there is a point in which adding an extra sensitivity decreases dramatically the

value of the minimum eigenvalue. The similarities of the vectors start to have a noticeable impact on the estimation from this point. Therefore, a good strategy consists of not selecting all PV candidates available but around 60% of the total available measures to get a rectangular shape for the sensitivity matrix and achieve accurate estimations.

3.4.3 Impact of the measurement vector time frame on estimations

Choosing the right time window is critical for the method to get accurate estimations. When more time points are included in the estimation, it is expected to increase the overall accuracy since more voltage increments due to PV injection are included. Figure 15 shows the average accuracy for 20 different single PV case scenarios from 11am to 1pm.

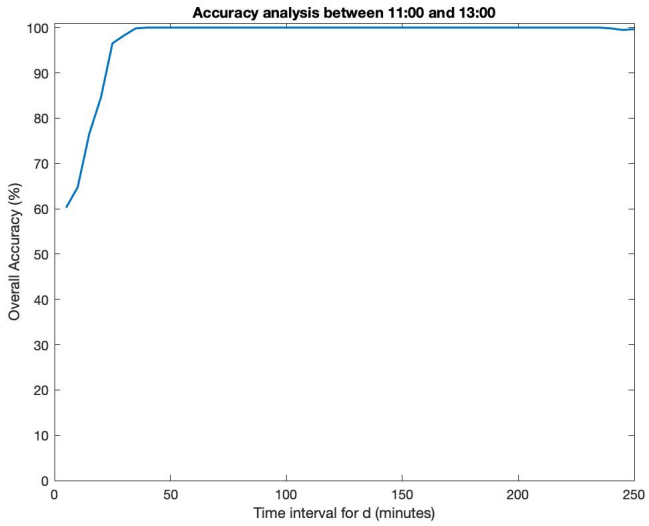


Figure 15 – Average accuracy for 20 different PV location in IEEE 13-bus feeder based on time frame for vector d from 11AM to 1PM

Unless the time window includes major power injections of PV, the precision of the estimation may be decreased compared to a single point in time window that captures a particular small injection. This situation specially happens when sun rises or sets. However, we achieve really precise estimations during intermittent cloudy days at noon because the voltage experiments many changes due to the difference in power injection. This points in time and time windows should be the target to make estimations. Figure 16 shows the impact of different time frames for the measurement vector.

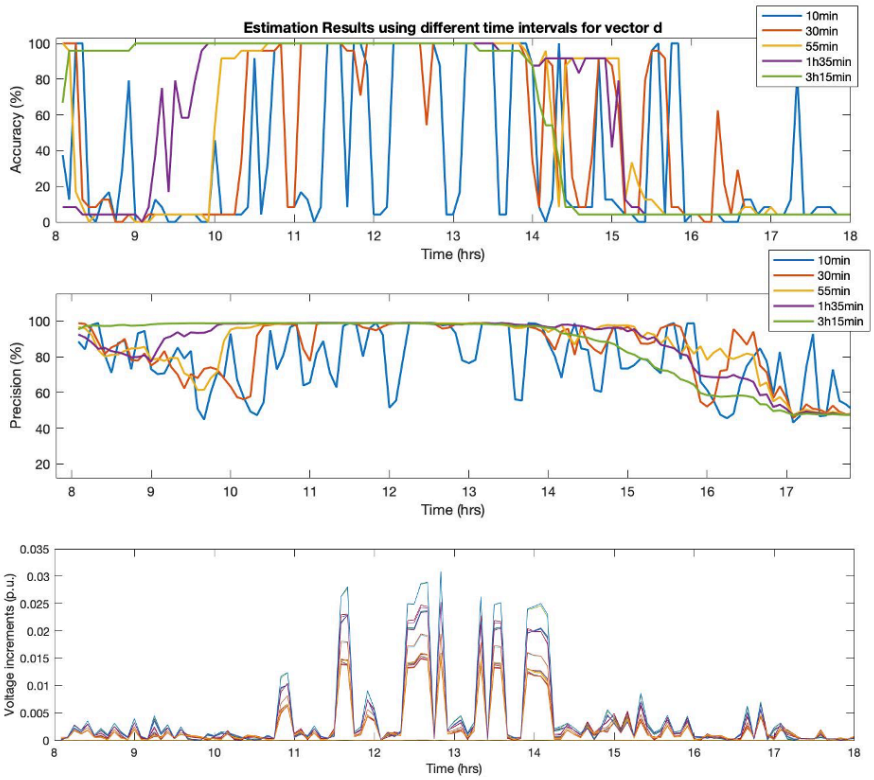


Figure 16 – Estimation accuracy results for different time intervals.

As can be observed, low time intervals are more sensitive to punctual voltage increments, while longer time periods get high accuracies for longer periods of time. Higher precisions usually go together with high accuracies, which means there exist no biases. Confidence is not affected by the range to evaluate $\boldsymbol{d}$. It depends upon whether the PV location to estimate is included or not and the number of degrees of freedom.

3.5 Multiple PV

Based on the sensitivity's linearity assumption, it is also possible to estimate multiple PV systems locations. This means that the resulting voltage increase of 2 PV systems will be the sum of each sensitivity proportional to the amount of power that each PV injects. Likewise, the resulting sensitivity vector of a 3-phase PV system will be the result of the corresponding sum of the 3 1-phase PV node location sensitivity vectors that are associated with that bus. This relation was verified by looking at the sensitivities of single and multiple PV case scenarios. Figure 17 shows compares 3 monitored nodes sensitivities components of single and multiple PV case scenarios.

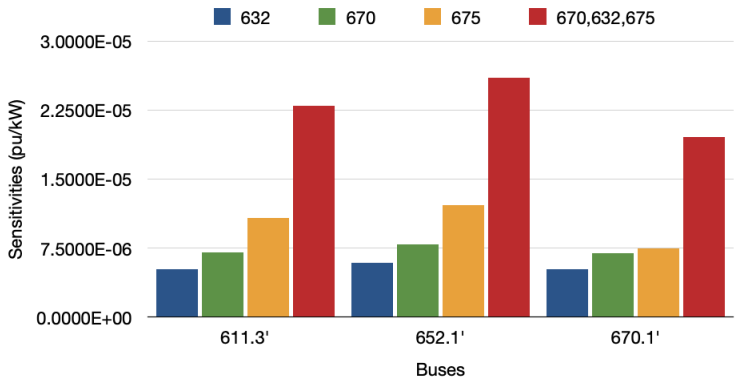


Figure 17 – Sensitivities of buses 611C, 652A, and 670A in terms of active power injections at buses 632, 670, 675, and 632, 670 and 675 at the same time.

Note that sum of the sensitivity components yields approximately to the multiple PV case sensitivity. This small error is introduced by the non-linear nature of the distribution systems. However, the model proposed is still valid because these are minimal, as can be observed on Figure 18.

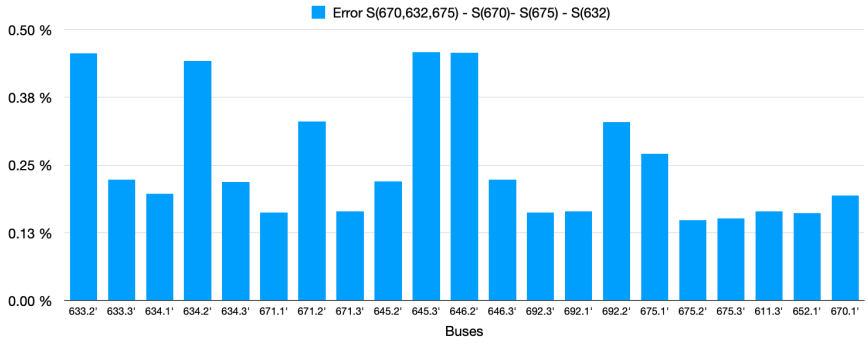


Figure 18 – Percentual sensitivity error for 22 buses of the system when we compute the difference of the multiple PV sensitivity.

The highest error component did not reach 0.5%. This sensitivity error is therefore considered neglectable. As result, we still can still consider a linear model under multiple PV case scenarios. Thus, we have a decoupled impact on the estimation, since for 2 different PV injection A and B it is verified that:

$$\hat{\mathbf{x}}^{A+B} = (\mathbf{S}^{\mathrm{T}}\mathbf{S})^{-1}\mathbf{S}^{\mathrm{T}}\Delta\boldsymbol{V}^{\boldsymbol{A+B}} = (\mathbf{S}^{\mathrm{T}}\mathbf{S})^{-1}\mathbf{S}^{\mathrm{T}}(\Delta\boldsymbol{V}^{\boldsymbol{A}} + \Delta\boldsymbol{V}^{\boldsymbol{B}}) =$$

$$= (\mathbf{S}^{\mathrm{T}}\mathbf{S})^{-1}\mathbf{S}^{\mathrm{T}}\Delta\boldsymbol{V}^{\boldsymbol{A}} + (\mathbf{S}^{\mathrm{T}}\mathbf{S})^{-1}\mathbf{S}^{\mathrm{T}}\Delta\boldsymbol{V}^{\boldsymbol{B}} = \hat{\mathbf{x}}^{A} + \hat{\mathbf{x}}^{B}$$

In the ideal case, we shall estimate non-zero injections for those PV candidates that hold a PV while the rest of components should be close to zero.

Figure 20 and Figure 19 shows some examples of multiple PV estimations in terms of heatmaps and spider plots for the 2 feeders under study.

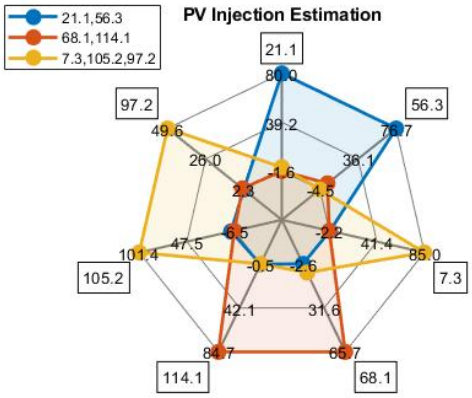


Figure 19 – Estimation results for 3 multiple PV case scenarios in the IEEE 123-bus test feeder: 2 PVs at 21A and 56C, 2 PVs at 68A, 114A and 3 PVs at 7C, 105B and 97B for the time frame from 11:35am to 11:40am.

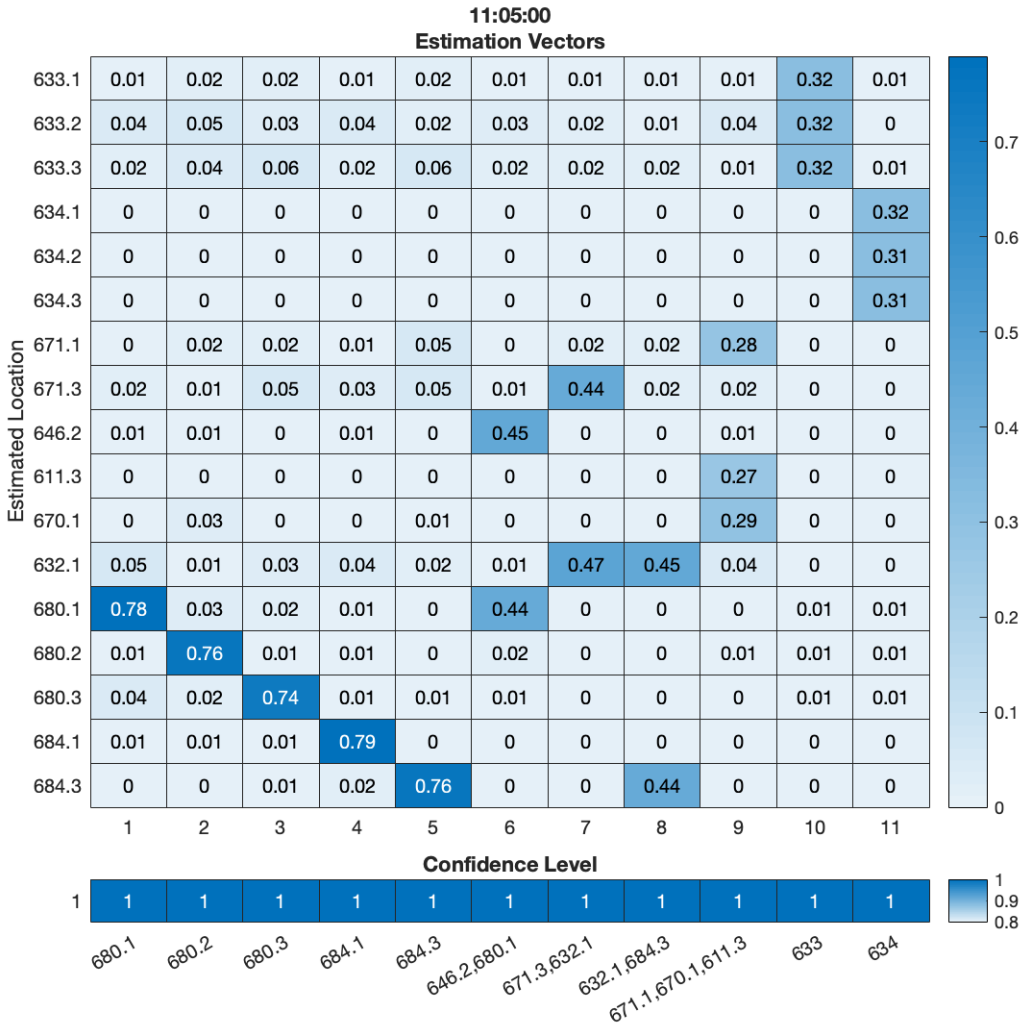


Figure 20 - Estimation results in principal components for single 1-phase PV, multiple 1-phase PV and single 3-phase PV system case scenarios in the IEEE 13-bus feeder

As can be observed, the method successfully predicts the location of multiple PV systems as well.

CHAPTER 4. IMPACT OF DISCRETE VOLTAGE REGULATION EQUIPMENT

The main limitation of the formulation presented in Chapter 3 is that it only admits changes in voltages due to PV or demand load. When a capacitor switches or tap changes position, the system voltages change significantly and no longer fall within the PV sensitivity plane defined by $\boldsymbol{S}$**.** In the following sections, the impact of these devices is analyzed, and the method is further extended to support the estimation of PV injection in the presence of VRE actions. This chapter presents a methodology to include the impact of discrete voltage regulation equipment on the estimations.

4.1 Capacitor Impacts

When the voltage magnitude data stream includes the effect of capacitor switching, the change in voltage will not result in an accurate estimation because it is caused by the injection of reactive power, and this sensitivity vector is not included in the $\boldsymbol{S}$ matrix. Typical discrete changes in voltage due to a capacitor can be observed in Figure 21. The objective is not only to detect when these discontinuities happen to prevent imprecise estimations, but to also try to achieve accurate PV location estimation at those time points.

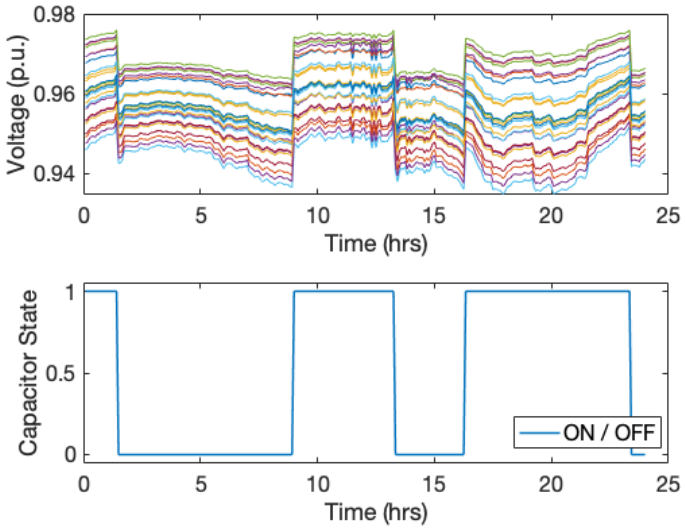


Figure 21 – Typical change in voltage due to a switching capacitor.

In order to incorporate the impact of capacitors switching in the method, one has to: a) estimate the impact on the active power sensitivities, b) compute the reactive power sensitivity and c) integrate the capacitor sensitivity as an additional column in the S matrix.

When a capacitor is connected to grid, the inductance matrix of the power system changes. Therefore, all the sensitivities are expected to change. The first task carried out was to check whether this change would impact greatly the PV sensitivities, which are critical for the estimation algorithm. The capacitor was connected and disconnected using OpenDSS, followed by the calculation of the active power sensitivities. For a 500 kvar 3-phase capacitor, the PV sensitivity difference between states was under 2%, and the difference of most of sensitivity components was minimal. To illustrate this, the S matrix

formed by 3-phase PV location candidates was calculated both with the capacitor connected and disconnected. Figure 22 shows the sensitivity differences between the 2 matrices. Note that the highest difference remains below 1.4%.

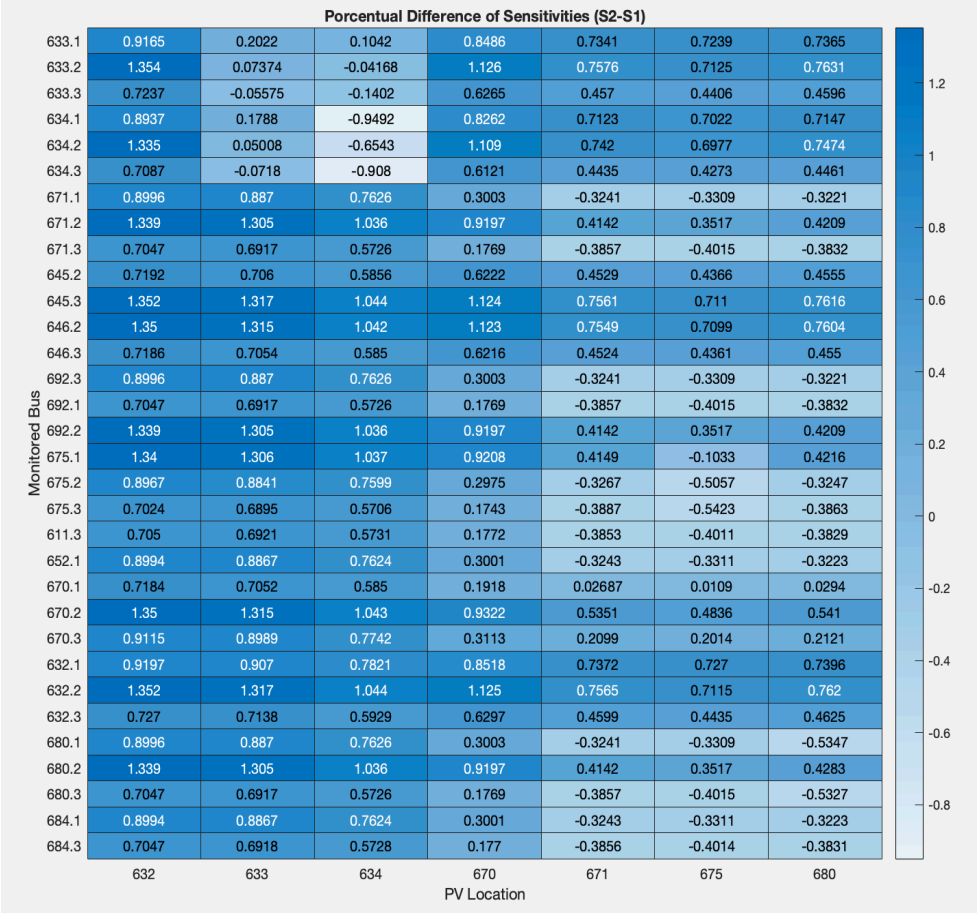


Figure 22 – Heatmap of the percentual difference of sensitivity matrices with and without capacitor bank.

In addition, when the estimations are computed using either of both S matrices, the difference between the final results is minimal. Figure 23 shows an example of an estimation using both sensitivity matrices.

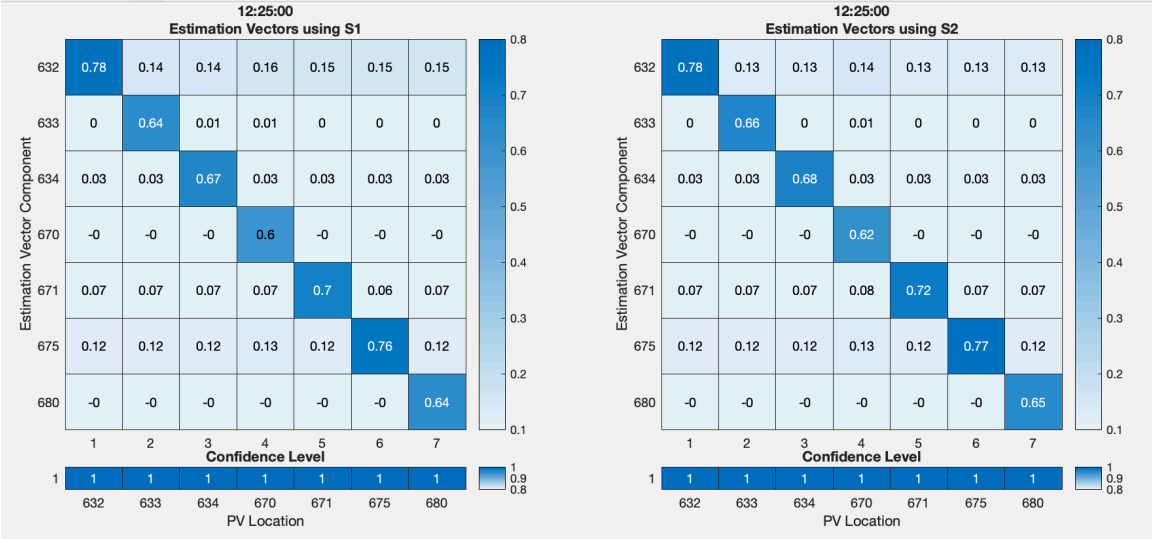


Figure 23 – Resulting Estimations using matrices S1 and S2 for a voltage measurement profile that includes the effect of the capacitor bank

This means that the state of the existing switching capacitor has minimal impact on the active power sensitivities of the distribution network, and the S matrix can be considered practically constant regardless of whether the capacitor status. However, when the capacitor switches between states, a discontinuity is introduced in the voltage stream. Figure 24 the two sensitivity planes, which are largely parallel to each other. The increment of voltage due to a change in net demand will be contained in those planes. However, when the voltage steps out of the voltage control region, the capacitor changes its status resulting in a voltage increase – symbolized as a red vector that is not necessarily vertical – that is not contained in the sensitivity plane defined by S. As result, the measurement vector cannot be expressed in terms of the sensitivity vectors and the estimation becomes inaccurate.

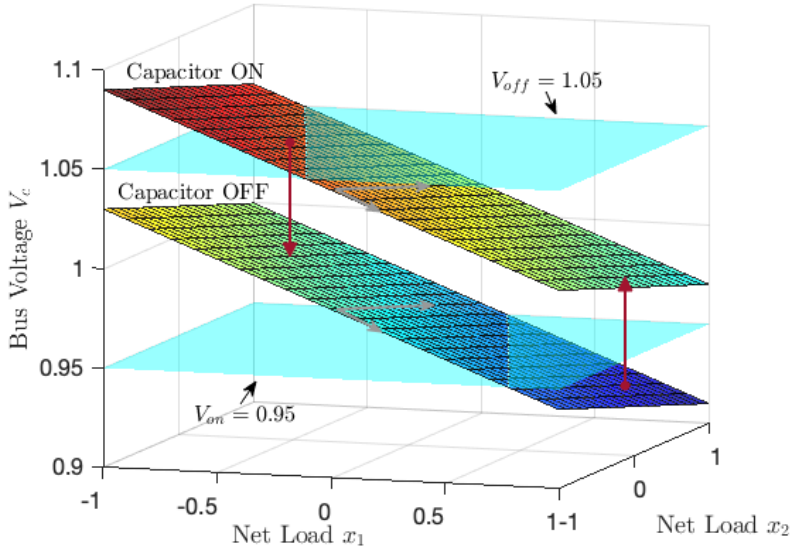


Figure 24 – Change of sensitivity planes due to capacitor state change.

Hence, our goal is to identify those points in time where the capacitor switches between states, while preserving the same sensitivity vectors that form the $\boldsymbol{S}$ matrix. In order to achieve this, a modification of the sensitivity matrix is introduced:

$$\boldsymbol{S} = \left[\boldsymbol{s_{PV}} \quad \boldsymbol{s_{Cap}}\right] \tag{4.1}$$

Here, $\boldsymbol{s_{PV}}$ corresponds to the set of columns of active power sensitivities that were already included, and s_{Cap} corresponds to the sensitivities associated with voltage changes in terms of reactive power injected by the capacitor. The sensitivities $\boldsymbol{s_{Cap}}$ can be computed by running two power flows: first with the capacitor disabled and then manually changing the state in OpenDSS, allowing us to compute the voltage difference due to the capacitor bank.

As an example, Figure 25 shows the voltage sensitivities for a 3-phase capacitor bank at bus 680. The larger the circles in the Figure, the higher the sensitivity.

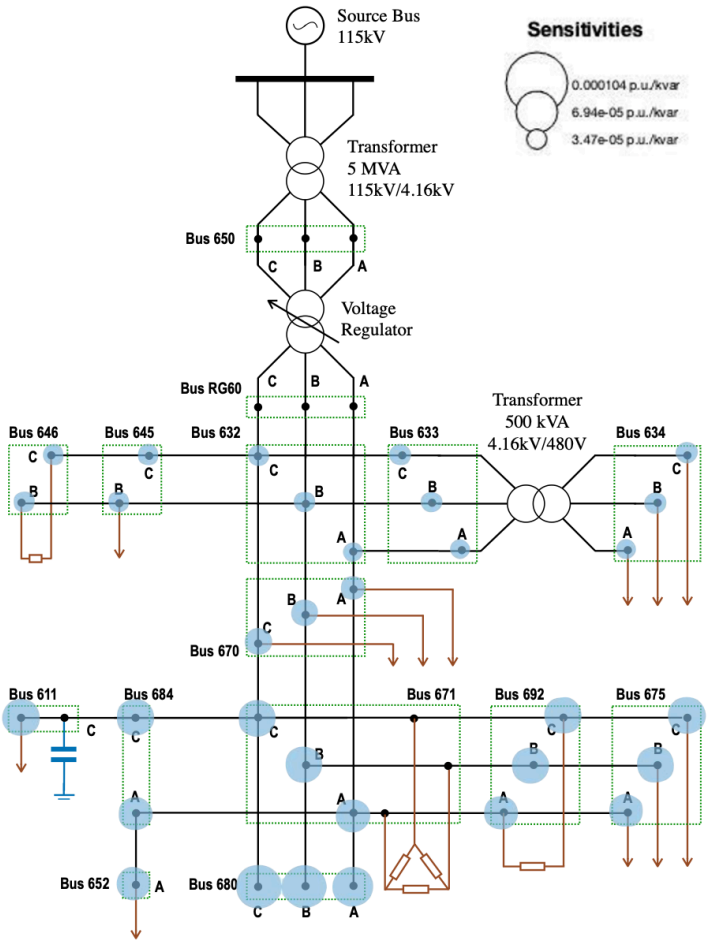


Figure 25 – Voltage increase due to the switching capacitor at bus 680.

Since the size of the capacitor is known, these sensitivities can be adjusted in such way that the estimation will indicate the percentage of a capacitor state change:

$$s_{i,Cap} = \frac{\Delta V_i}{Q_{Cap}} \tag{4.2}$$

such that when a voltage increase due to capacitor occurs, the estimation components associated to capacitor actions shall be close to 100%.

4.2 Voltage Regulator Impacts

Similarly, voltage regulators will cause a discontinuity in the voltage profile when the tap position changes, leading to a sensitivity hyperplane that practically has the same slope as that of the previous tap position. Figure 26 shows both planes when a change of voltage due to PV and tap position change occurs. Note that the resulting measurement vector (in blue) does not belong to the sensitivity due to the VRE component.

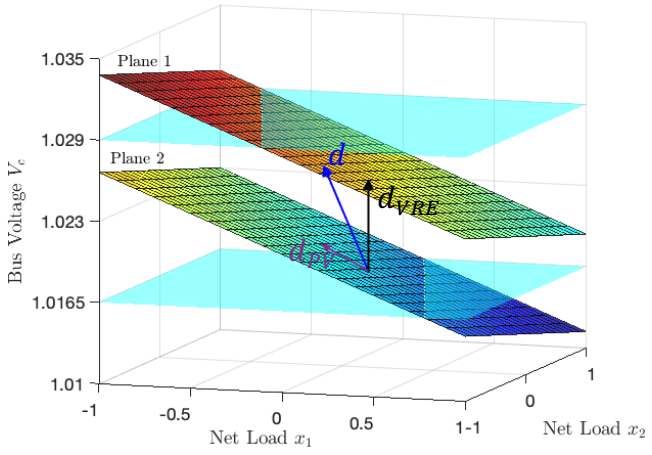


Figure 26 – Change of sensitivity planes as a consequence of a tap change.

In order to estimate both the tap change and the PV injection, a sensitivity associated to the voltage regulator action needs to be included in the $\boldsymbol{S}$ matrix. The VRE sensitivities are calculated by changing the taps on OpenDSS and subtracting the voltage recorded from both power flows. A tap change produces a homogenous increment in voltage of the same-phase node that comes after the regulator. Figure 27 shows the voltage increments of a tap change for the bus 632A voltage regulator.

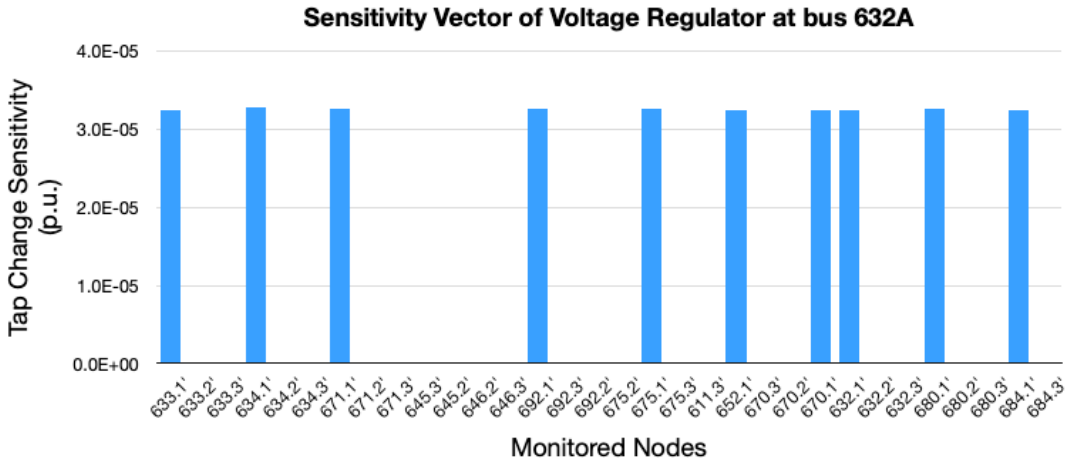


Figure 27 – Impact of voltage regulator at bus 632A in IEEE 13 feeder

Once all VRE sensitivities have been determined, the modified S matrix is formed as:

$$\boldsymbol{S} = \left[\boldsymbol{s_{PV}} \quad \boldsymbol{s_{taps}}\right] \tag{4.3}$$

where the s_{taps} vector is calculated as:

$$s_{i,tap} = V_{i,tap+1} - V_{i,tap} \tag{4.4}$$

Note that the estimation is done in terms of percent of tap step. For instance, if the tap changes 2 positions we ideally should get a result close to 200% for the voltage regulator component in the estimation vector.

4.3 Estimation in Feeder with Interacting Voltage Regulating Devices

The methodologies developed in sections 4.1 and 4.2 are based on the use of a matrix of voltage sensitivities with respect to node power injections $\boldsymbol{S}$. However, in the presence of VRE, the impact of VRE actions and the PV power injection both may produce changes in voltage magnitude in the circuit nodes that result in sensitivity columns that are not linearly independent vectors. To illustrate this, let us consider Figure 28, where a voltage change that is due to a VRE action can be expressed in terms of the PV sensitivity column vectors.

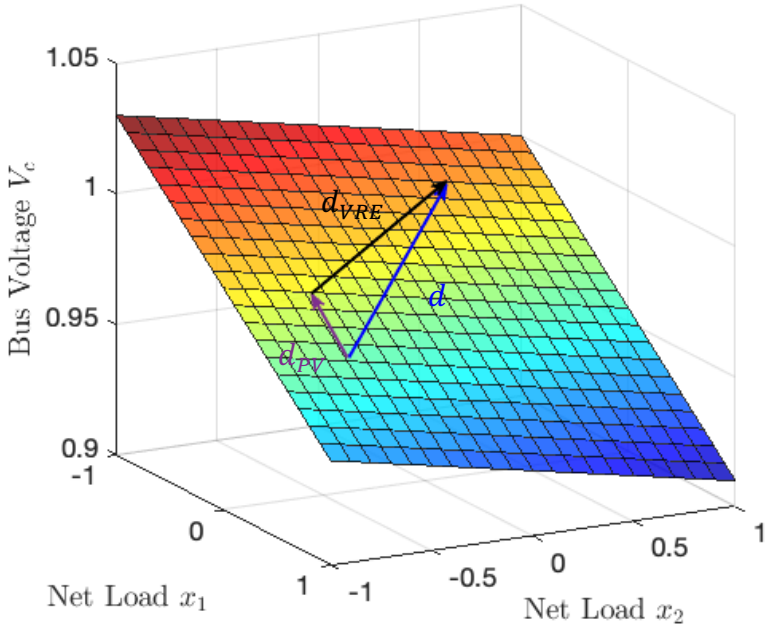


Figure 28 – VRE impact that is linearly dependent on PV sensitivities.

In such case, the Gramian of $\boldsymbol{S}$ becomes singular and $(\boldsymbol{S}^T\boldsymbol{S})^{-1}$ cannot be computed. Therefore, an approach must be developed that can estimate PV injections, tap changes, and switching capacitor actions regardless of the structure of VRE sensitivities.

The method extends the principles used for PV location and injection estimation by first determining whether a VRE actions is present in the voltage magnitude measurement data stream. Let us define $\widehat{\boldsymbol{d}}_{VRE}$ as the estimation of the changes in voltage due to VRE actions in a distribution circuit. The correct estimation of the PV injections would discount the effect due to the VRE in changes in voltages $\boldsymbol{d}$:

$$\hat{\boldsymbol{x}} = (\boldsymbol{S}^T\boldsymbol{S})^{-1}\boldsymbol{S}^T(\boldsymbol{d} - \hat{\boldsymbol{d}}_{VRE}) \tag{4.5}$$

The changes in voltage occur in fixed amounts, proportional to the number of taps changed. It is possible to obtain an estimation of a voltage changes due to VRE, $\boldsymbol{d}_{VRE}$ by simulating VRE actions in the distribution circuit and performing an estimation for the resulting voltage changes. Let us denote the result of this estimation by $\hat{\boldsymbol{x}}_{VRE}$:

$$\hat{\boldsymbol{x}}_{VRE} = (\boldsymbol{S}^T\boldsymbol{S})^{-1}\boldsymbol{S}^T\boldsymbol{d}_{VRE} \tag{4.6}$$

This particular estimation vector corresponds to a single VRE device. A matrix $\boldsymbol{X}_{VRE}$ can be formed when considering all the VRE devices in the circuit:

$$\boldsymbol{X}_{VRE} = (\boldsymbol{S}^T\boldsymbol{S})^{-1}\boldsymbol{S}^T\boldsymbol{D}_{VRE} \tag{4.7}$$

where $\boldsymbol{X}_{VRE} = [\hat{\boldsymbol{x}}_{VRE_1}, \ \ldots, \hat{\boldsymbol{x}}_{VRE_K}]$, $\boldsymbol{D}_{VRE} = [\boldsymbol{d}_{VRE_1}, \ \ldots, \boldsymbol{d}_{VRE_K}]$ and K is the number of total VRE devices considered. We have that each vector $\hat{\boldsymbol{x}}_{VRE}$ in $\boldsymbol{X}_{VRE}$ is the expected footprint that a tap change will leave in the estimation. Once $\boldsymbol{X}_{VRE}$ has been determined, it can be used to determine the presence of tap change actions, if the resulting estimation vector is similar to $\hat{\boldsymbol{x}}_{VRE}$. When a change in voltage contains the impact of VRE and PV for a given point in time, the resulting estimation vector will contain components associated with both the PV location and the expected estimation $\hat{\boldsymbol{x}}_{VRE}$:

$$\boldsymbol{x}_{PV+VRE} = (\boldsymbol{S}^T\boldsymbol{S})^{-1}\boldsymbol{S}^T\boldsymbol{d} \tag{4.8}$$

The matrix $\boldsymbol{X}_{VRE}$ can now be used to determine whether any VRE action took place by performing a second estimation on the resulting estimation vector:

$$\boldsymbol{v} = ((\boldsymbol{X}_{VRE}^T\boldsymbol{X}_{VRE})^{-1}\boldsymbol{X}_{VRE}^T) \cdot \boldsymbol{x}_{PV+VRE} \tag{4.9}$$

The resulting vector v will provide non-zeros values for those components associated with the VRE devices that operated at that specific point in time. For example, if 2 step changes occurred for a certain voltage regulator, the resulting $\boldsymbol{v}$ component may be 2.03. For those devices that did not take action at that point in time, a value close to 0 will appear. A non-linear filter ϕ needs to be applied to remove the values close to zero and to obtain integer components from $\boldsymbol{v}$. For illustration purposes, let's consider the following example. Assume that there are 3 voltage regulators in a distribution circuit, and the results of our second estimation is: $\boldsymbol{v} = [0.04 \quad 0.98 \quad -0.01]^T$, then the filter $\phi(\boldsymbol{v}) = [0\ 1\ 0]^T$, which corresponds to the actual tap changes that occurred at that point in time. The filter should be chosen based on what provides better estimation results. Some examples of filters are shown in Figure 29.

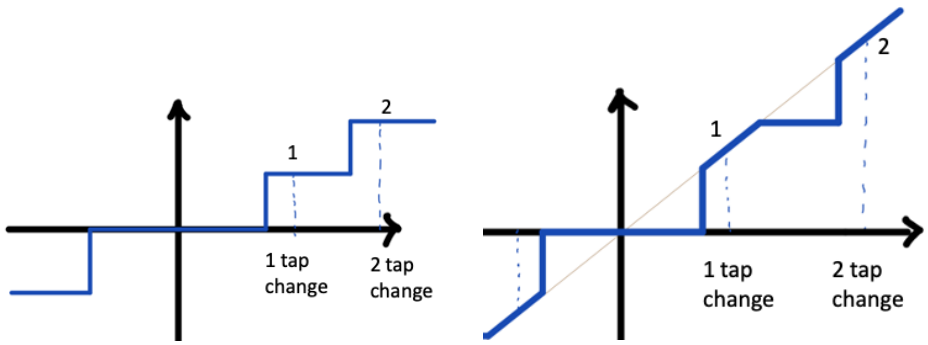


Figure 29 – Examples of tap changes filter

The filter may have a significant impact on the results, as it a way to calibrate how much of the expected VRE voltage increment should we subtract from the measurement vector. For example, Figure 30 shows the implementation using the stairs-shape filter shown on Figure 29 left. If the filter is changed to the right one, the resulting estimation no longer has the some of the component off that we were previously getting, as can be observed in Figure 31. The filter selection is important to achieve high accuracy results.

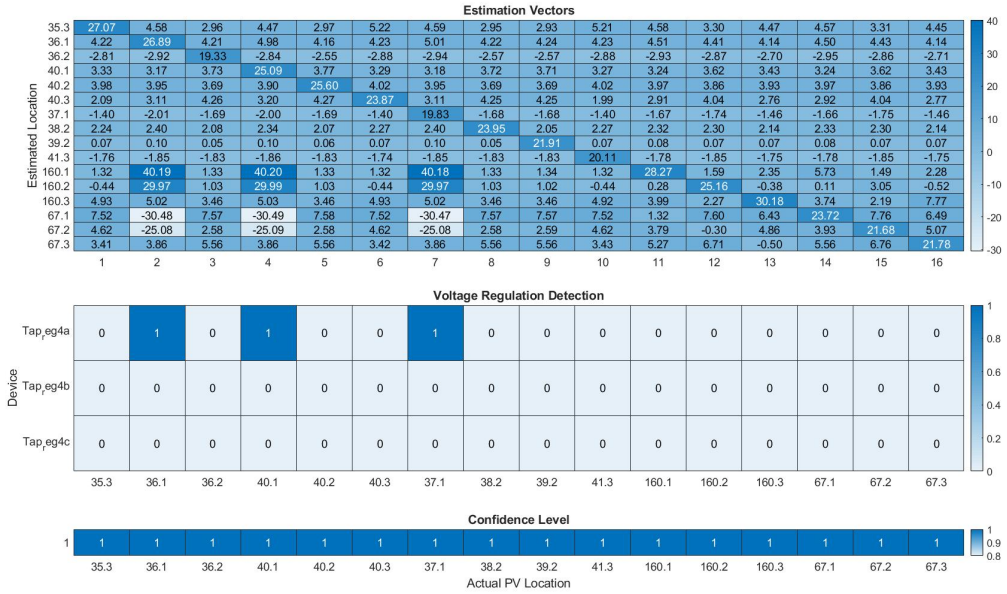


Figure 30 – Estimation results using filter from Figure 29 (left).

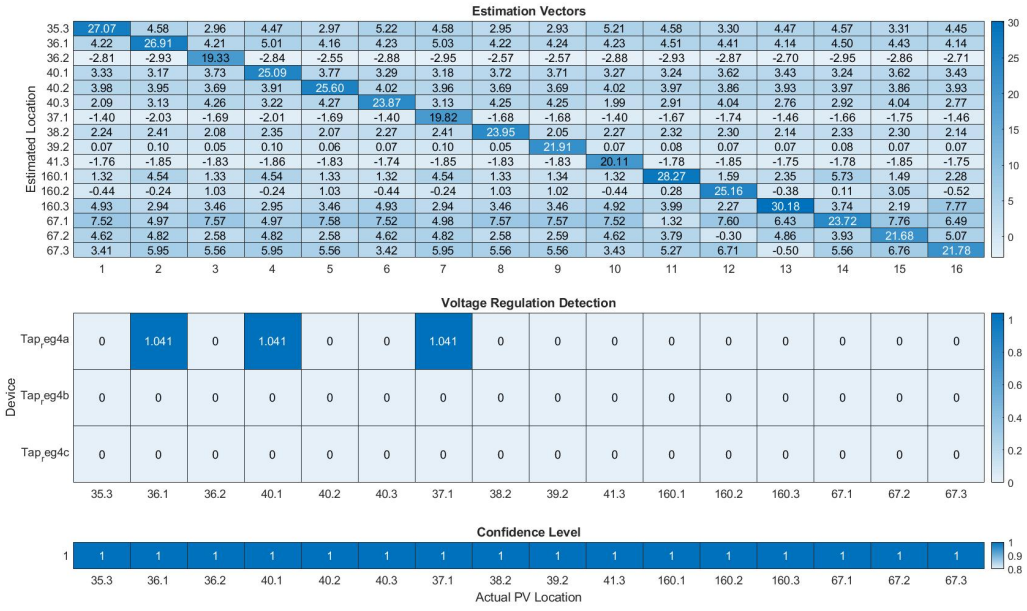


Figure 31 – Estimation results using filter from Figure 29 (right).

Once a VRE action is detected, the impact on the voltage is computed by multiplying by $\boldsymbol{X}_{VRE}$, which leads to the estimation components of tap change, and finally by $\boldsymbol{S}$, which leads to the estimated voltage increase due to that VRE action:

$$\hat{d}_{VRE} = \boldsymbol{S} \cdot \boldsymbol{X}_{VRE} \cdot \phi(\boldsymbol{v}) \tag{4.10}$$

With the subtraction of the impact of VRE on voltage increments it is not only possible to predict the location of PV systems, but also the action taken by controlling devices.

Putting together equations (4.7), (4.8), (4.9), and (4.10) in (4.5) the formulation of the method now becomes:

$$\hat{\boldsymbol{x}} = (\boldsymbol{S}^T\boldsymbol{S})^{-1}\boldsymbol{S}^T\left(\boldsymbol{d} - \boldsymbol{S} \cdot \boldsymbol{X}_{VRE} \cdot \phi\left(((\boldsymbol{X}_{VRE}^T\boldsymbol{X}_{VRE})^{-1}\boldsymbol{X}_{VRE}^T)(\boldsymbol{S}^T\boldsymbol{S})^{-1}\boldsymbol{S}^T\boldsymbol{d}\right)\right) \tag{4.11}$$

where:

- $\boldsymbol{S}$ is the sensitivity matrix. It has N rows and L columns.
- $\boldsymbol{d}$ is the measurement vector: $[\boldsymbol{d}_1, \dots, \boldsymbol{d}_N]^T$.
- $\boldsymbol{X}_{VRE}$ includes all the estimation vectors due to VRE devices. It has L rows and K columns.
- $\boldsymbol{v}$ is the vector of estimated VRE actions. Hence its dimension is equal to the number of total VRE devices in the circuit: $[\boldsymbol{v}_1, \dots, \boldsymbol{v}_K]^T$.

This is the general formulation that is used to estimate PV injections for a given point in time considering VRE actions. The number of tap changes can also be predicted by looking at the term $\phi\left(((\boldsymbol{X}_{VRE}^{T}\boldsymbol{X}_{VRE})^{-1}\boldsymbol{X}_{VRE}^{T})(\boldsymbol{S}^{T}\boldsymbol{S})^{-1}\boldsymbol{S}^{T}\boldsymbol{d}\right)$.

4.3.1 Numerical Results

The IEEE 123-bus system presents switches that have been configured to preserve a radial topology. It is assumed that voltage magnitude measurements are obtained from Intellirupters, which cover 8% of the per phase voltage magnitudes. The circuit topology and Intellirupters location are presented in Figure 32.

To enable the algorithm to work with less available measurements, the number of PV location candidates must be lower to avoid singularities in the computations. In the results that follow below, different PV case scenarios are simulated to test the methodologies presented to address both linear dependent and independent impact of VRE. The time resolution selected for the QSTS simulations is 300 seconds.

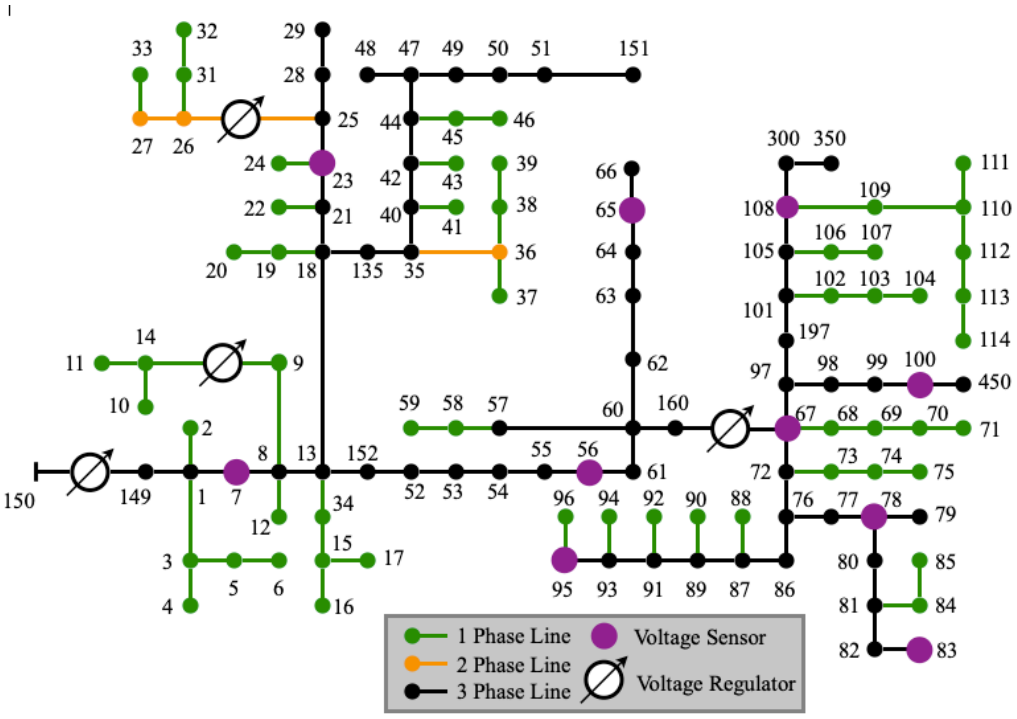


Figure 32 – Diagram of IEEE 123-bus system with limited voltage measurements.

The first experiment was run considering single 1-phase 100 kW PV at buses 23, 197 and 83. In order to get an accurate estimation, it is necessary to look at points in time that are partly cloudy so that the node experiences an increase of PV injection. Based on the selected PV profile, the time window 12:25pm to 12:30pm is appropriate to estimate a PV injection. Figure 33 shows the results of an estimation for this time frame. Out of the 9 simulations, only the one with PV at bus 23.1 experienced a tap change in voltage regulator 67A. As can be observed, the method is able to capture both the change in the PV injection (approximately 25 kW at that point in time) and the tap change.

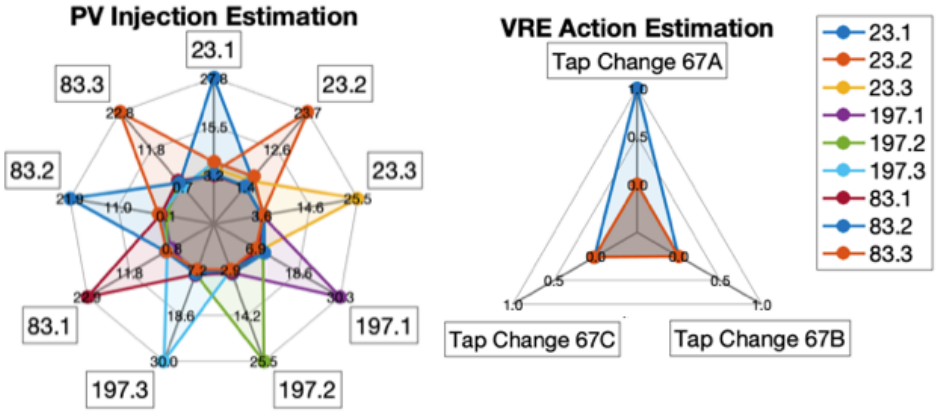


Figure 33 – PV and VRE Estimation results for simulations from 12:25pm to 12:30pm of 12 scenarios with a single PV

Now let us consider the particular case when the PV and VRE impacts are linearly dependent. This conflict causes singularities in the gramian of $\boldsymbol{S}$ and hence it is necessary to use the formulation presented in (4.11). To represent this particular case, some buses close to bus 67 have been chosen as potential candidates. Let us consider buses 67, 62 and 68 as potential candidates. When $\boldsymbol{X}_{VRE}$ is calculated, it is verified that the voltage regulator sensitivities can be expressed in terms of the sensitivity vectors associated to PV located at bus 62 and 67. Figure 34 shows a bar chart of the vectors that compound $\boldsymbol{X}_{VRE}$. It can be observed that the impact of VRE affects the estimation at buses 67 and 62, but has little significant impact on bus 68. In addition, when $\widehat{\boldsymbol{D}}_{VRE} = \boldsymbol{S} \cdot \boldsymbol{X}_{VRE}$ is computed, we find that $\widehat{\boldsymbol{D}}_{VRE} = \boldsymbol{D}_{VRE}$ and the confidence level is equal to 1. This indicates that the VRE sensitivity falls within the PV sensitivity hyperplane.

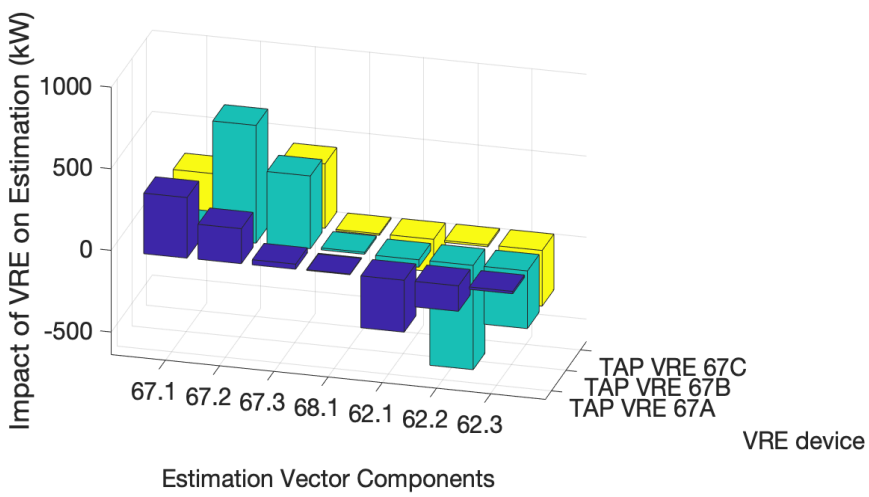


Figure 34 – Estimation of voltage increment due to VRE.

To test the accuracy of this methodology, several simulations were conducted placing PV systems at conflicting and non-conflicting nodes. These 7 simulations include a PV at each phase of buses 62, 67 and 68. The time window between 11:55am to 12:05pm had samples that included tap changes in addition to PV injection. Figure 35 shows the resulting estimation. The method provides high accuracy predictions of the PV injection in addition to the exact 3 tap changes that occurred in the simulations with a PV located at 62B, 67A and 68A. The experiments on the IEEE 13-bus were successful as well. This circuit includes impact of the 3 single-phase voltage regulators and the two switching capacitors located at buses 611C and 646B. Figure 36 (left) shows the resulting estimation for a point in time that included multiple tap changes, PV injection and switching capacitor impact.

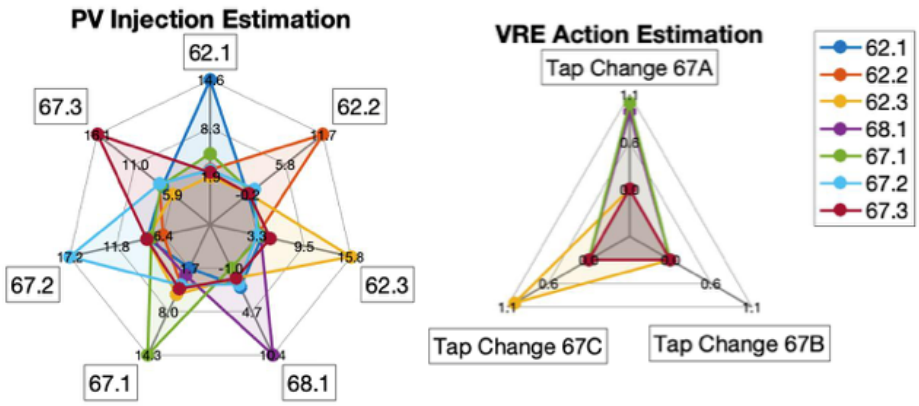


Figure 35 – PV and VRE Estimation for 7 simulations using (4.11).

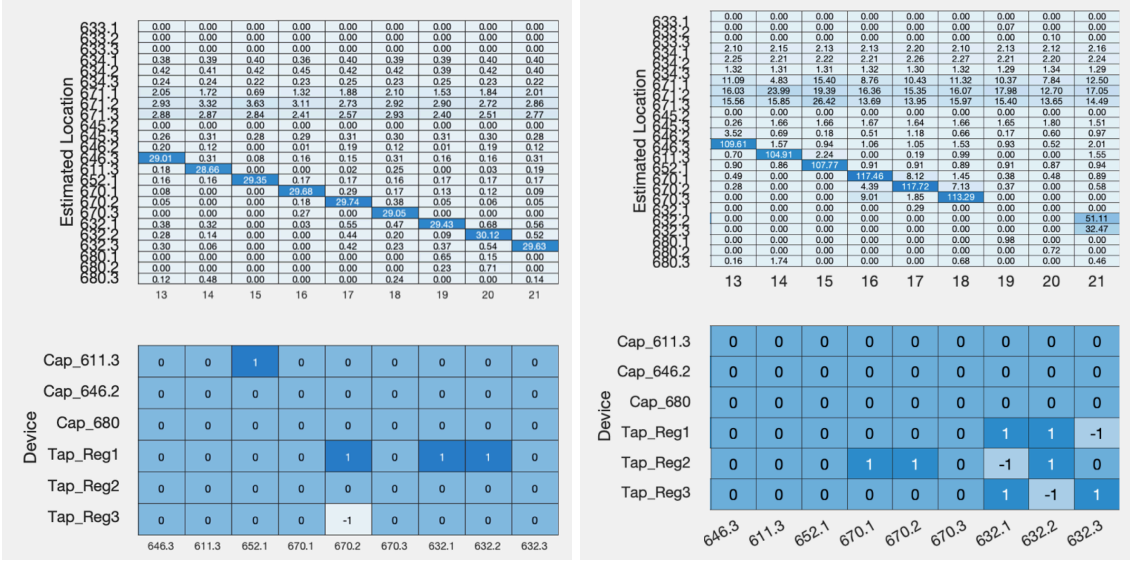


Figure 36 – Impact of PV injection on estimation using (4.11).

Nevertheless, there is one situation that the method will not be able to make accurate predictions: when there is a combination of PV injection that exceeds the estimation values of $\boldsymbol{X}_{VRE}$ Figure 36 (right). The reason behind this limitation resides in the fact that this

methodology assuming the VRE will have a unique linear combination of the PV sensitivity vectors that is fixed. When this combination is reached, the method will predict a tap change because the combination of multiple PV is producing a similar impact as the VRE in voltage magnitude terms.

CHAPTER 5. MISSING DATA, ERRORS AND LIMITATIONS

This chapter presents the analysis of the method performance when there is noise in the measurements, or some voltage data is missing. The limitations of the algorithm are also covered in the last section of the chapter.

5.1 Impacts of Errors in Measurements

So far, the method assumes that the measures provided by sensors do not contain noise. The inaccuracies introduced by noisy measures will impact the estimations. Since the estimation compares the sensitivity vector and measurement vector, the analysis presented in this section also applies to the inaccuracies of an imperfect model. This is particularly interesting since utilities usually do not have an exact model of the distribution circuit. Hence the calculated sensitivities may differ from the real ones up to a certain degree. This section studies how much noise in the measurements or inaccuracies in the sensitivities' calculation can admit the method to still provide accurate estimations.

The analysis is done with the IEEE 13-bus feeder and it is split into two parts. First, the theoretical implications are studied to see what the noise limit in theory is. Secondly, a set of estimations is run assuming different degrees of gaussian noise to observe the real impact on the predictions.

5.1.1 Theoretical Approach

In the theorical approach, it is assumed that the voltage increment is going it be exactly 100 times the sensitivity vector. That is, an ideal PV injection of 100kW is considered at multiple location scenarios:

$$\boldsymbol{D} = 100 \cdot \mathbf{S}$$
$$X = (\mathbf{S^T S})^{-1}\mathbf{S^T}\boldsymbol{D} \tag{5.1}$$

As result we expect $X = 100 \cdot I$, where I is identity matrix. This result let us compare the precision when gaussian noise in included in the measures. The evaluation of estimation is evaluated as follows.

1. Compute the standard deviation of noise from: $SNR = 20\log\left(\frac{1}{\sigma}\right)$, so $\sigma = e^{-\frac{SNR}{20}}$. According to [26], PMU data can have up to 45 dB of signal-to-noise ratio SNR (10% of error).

2. Compute $D = 100 \cdot S \cdot (1 + \nu)$, where $\nu \sim N(0, \sigma)$

3. Perform the estimation for that noise 100 times in order to homogenize results.

4. Calculate the average of estimated values and principal components.

5. The precision for that SNR will be $\sum \frac{diag(PC)}{n}$

Once the precision is calculate based on the SNR, we can compare the method accuracy based on the SNR and the number of degrees of freedom. Figure 37 presents three figures that show how the accuracy is impacted by these two variables. The discontinuous yellow line denotes a 45 dB SNR (10% or error in the sensor measurements).

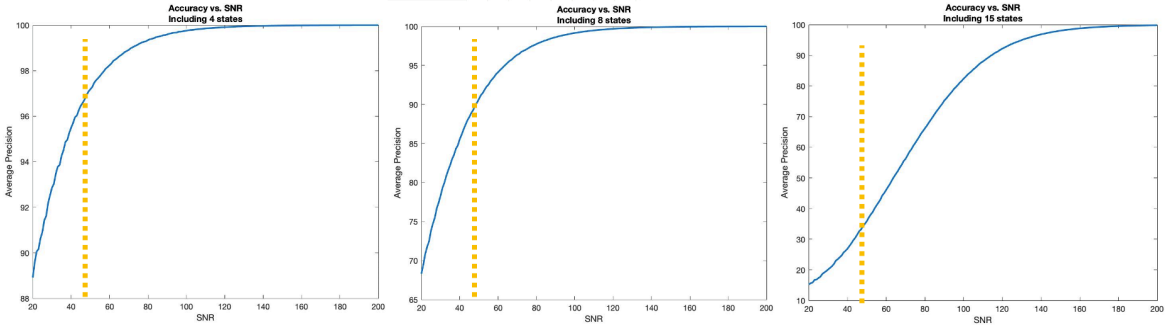


Figure 37 – Average Precision Curve including 4 (left), 8 (middle) and 15 (right) sensitivity vectors in the matrix S.

As can be observed, the more sensitivity vectors we include in the ***S*** matrix, the more likely is to misclassify the location of PV when we have noise in the data because we have more choices that might be more aligned with the resulting measurement vector. Similarly, the more noise in the measurement, the lower result precision, as we would expect. Figure 38 shows the accuracy in terms of both the SNR and the number of sensitivities on a 3D plot. A red plane is added to delimit the desired minimum acceptable precision (90%). The intersection between the resulting surface and this plane shows the trade-off in terms of SNR and degrees of freedom to achieve a theorical accuracy of 90%.

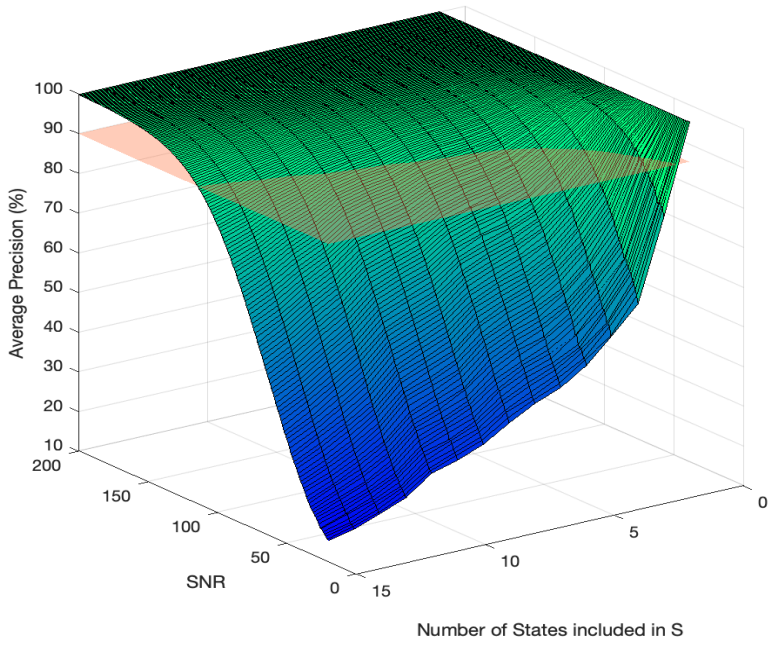


Figure 38 – Impact of SNR and the degrees of freedom on the theoretical method precision.

As result, adding more noise to the measurements forces us to reduce the number of PV locations candidates if we want to maintain a 90% of precision.

5.1.2 Numerical Results

To validate the previous analysis, gaussian noise was added to the measurements recorded by sensors for 8 different simulations in the IEEE 13-bus test feeder. The following figures show the estimations results for the time window 12:00pm to 12:05 as we increase the ratio of noise.

Figure 39 – Estimation results using the voltage difference between 12:00PM and 12:05PM with no noise.

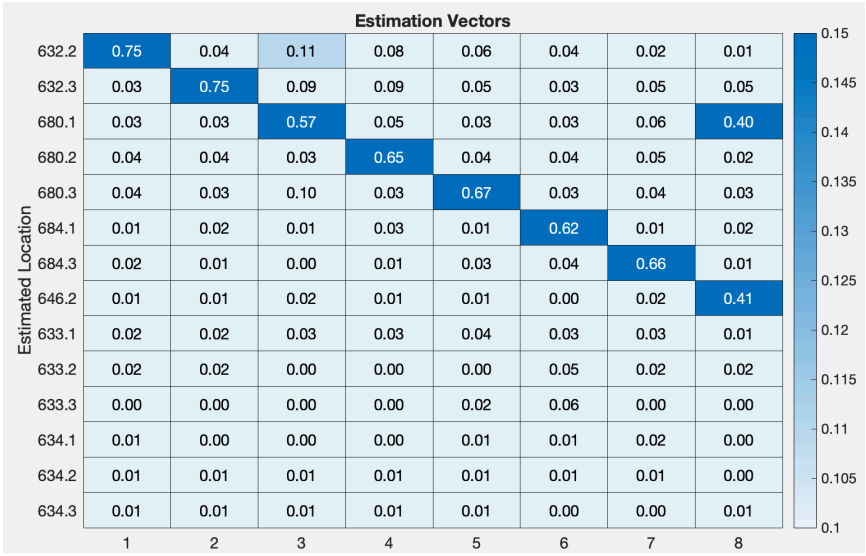


Figure 40 – Estimation results using the voltage difference between 12:00PM and 12:05PM with 80 dB SNR (1.8% noise)

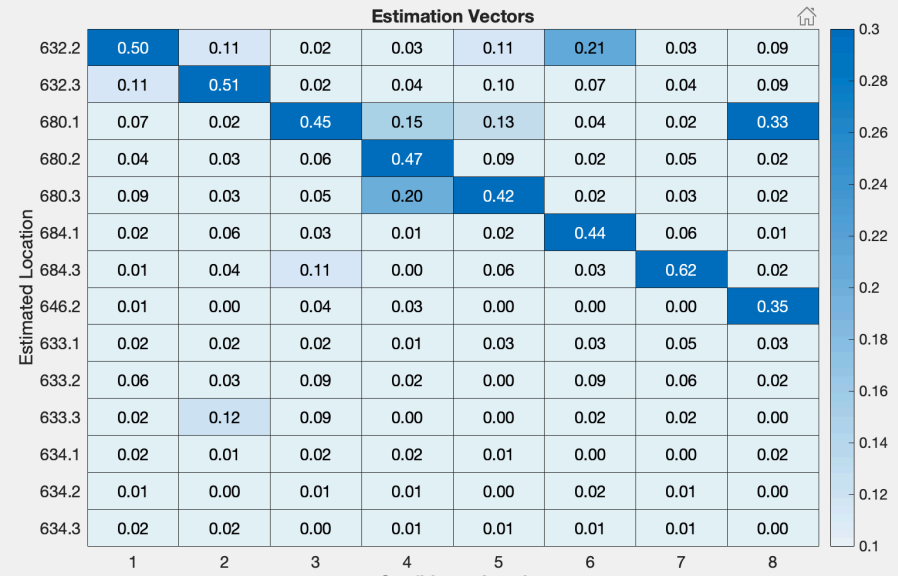


Figure 41 – Estimation results using the voltage difference between 12:00PM and 12:05PM with 60 dB SNR (5% noise)

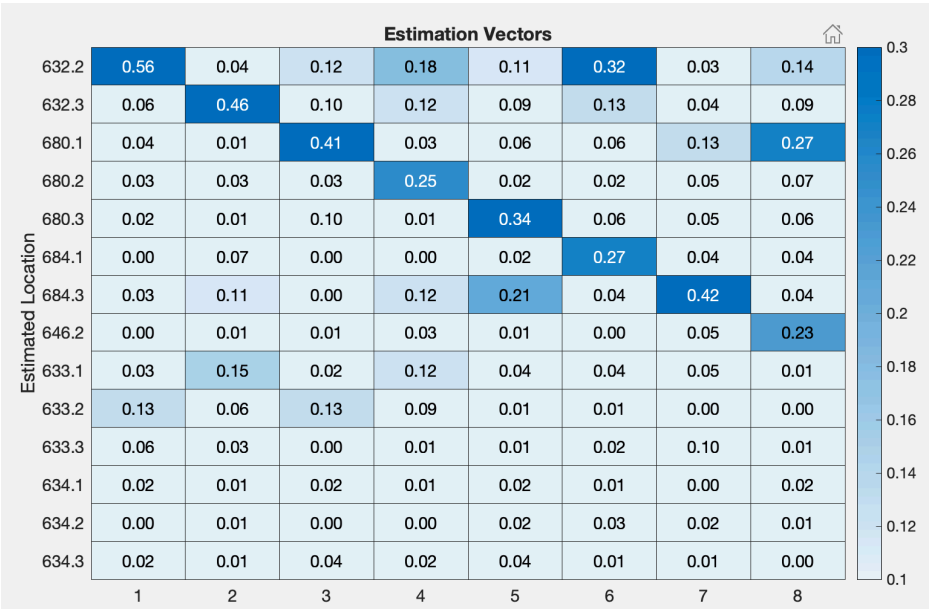


Figure 42 – Estimation results using the voltage difference between 12:00PM and 12:05PM with 45 dB SNR (10% noise).

As the noise increases, it is harder to get accurate estimations. While the simulation with 80dB of barely affects the estimated vectors (only simulation with PV at bus 632.2 is slightly affected), we see significant impact on the estimations that include 10% noise. In particular, the simulations with PV at buses 680B, 646B and 680A, 680C, and 684A experience significant biases towards incorrect nodes. To see the impact of degrees of freedom in the accuracy, sensitivity buses 633 and 634 were removed from the S matrix. Figure 43 shows the results of the new estimation while maintaining a SNR of 45dB.

Figure 43 – Estimation results using the voltage difference between 12:00PM and 12:05PM with 45 dB SNR (10% noise) with 6 extra degrees of freedom.

We see a clear improvement of accuracy with respect to Figure 42. With a lower set of PV location candidates, the accuracy improves as the inaccuracies introduced by noise cannot

lead to previous misclassifications (buses 633 and 634). The estimations remain accurate within the range of 90 to 100% if the sensitivity matrix keeps a dimension of 29x6 for the usual signal to noise ratios.

5.2 Algorithm Performance with Missing Measurements

So far, the method assumes complete availability of voltage measurements across the distribution feeder. However, most utilities do not have access to all the voltage per phase magnitudes and this impacts the estimation accuracy. This section covers the impact of reducing the number of degrees of freedom on the estimation due to the missing voltage data.

The methodology to deal with missing measurements consist of eliminating the row associated to the missing monitoring node from the **S** matrix. The estimation is performed as usual if the missing data does not produce conflicts in the **S** matrix: when excluding a component from sensitivity vector, the resulting vectors should still be independent. This not always holds, as there are particular cases in which two or more resulting sensitivity vectors can become linearly dependent and the inverse of $\mathbf{S^T S}$ is non-computable. These are very rare cases that may be solved with regularization techniques.

The main impact of excluding one row in the sensitivity matrix is that the number of possible PV candidates that we can predict is reduced to L-1, to be consistent with the degrees of freedom rule introduced in section 3.4.2. Let us take as an example the IEEE

123-bus circuit by assuming that only buses 13, 42, 52, 67 and 105 are being monitored. Buses 64, 135 and 83 are considered potential PV location candidates, leading to the rectangular S matrix that is shown in terms of a heatmap in Figure 44.

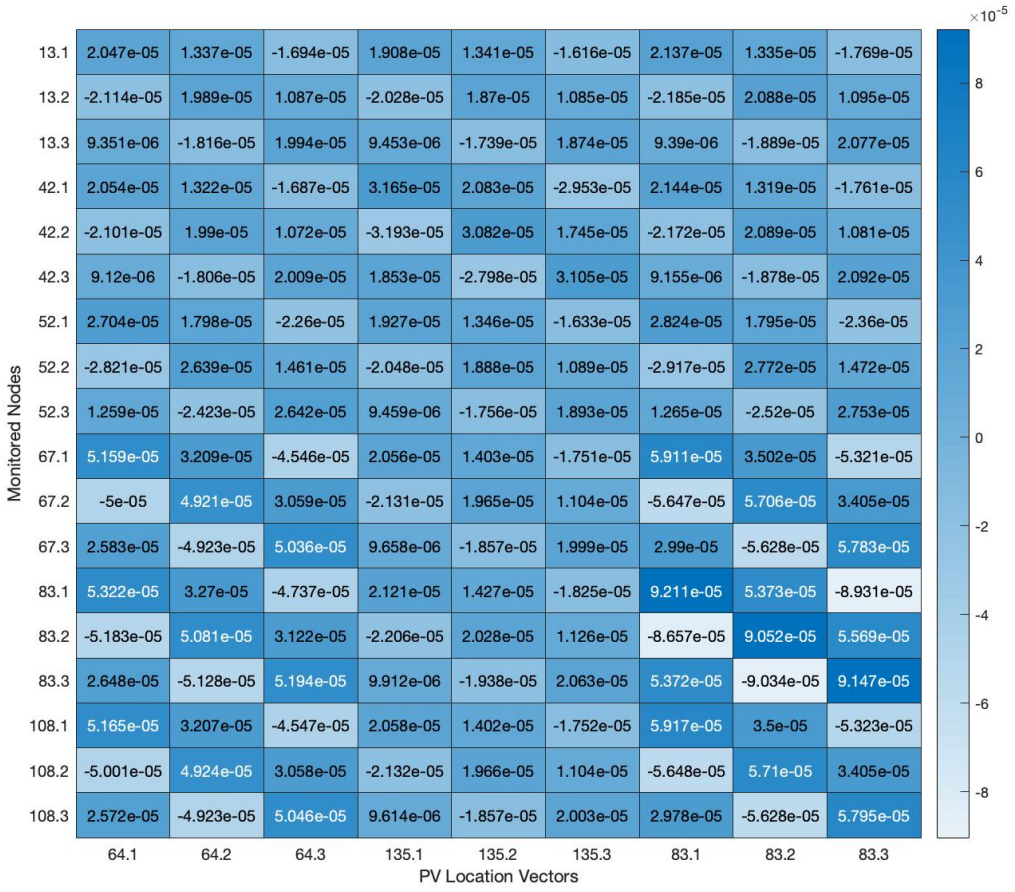


Figure 44 – Modified S Matrix. S We consider only nodes 13,42,52,67,83 and 105 are being monitored. The PV location candidates are each phase-node of buses 64, 135 and 83.

Note that this matrix is formed by sensitivity vectors that are not linearly dependent on each other. The difference between making an estimation with the complete S matrix and the modified one can be observed if we compare Figure 45 and Figure 46.

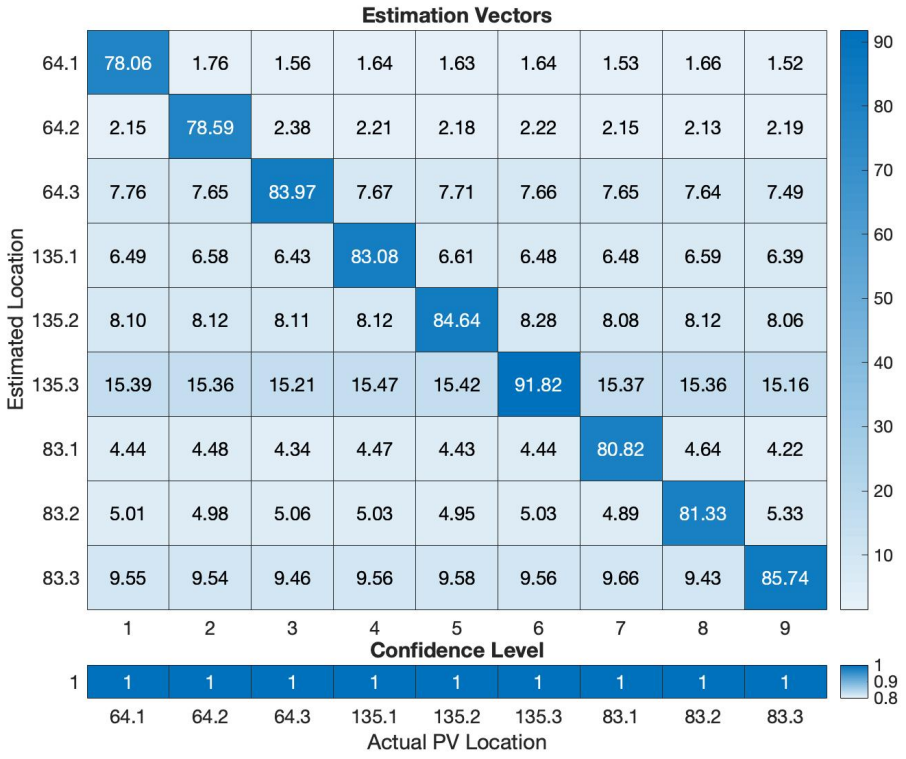


Figure 45 – Estimation results for 9 different simulations within the time window 12:25 to 12:30pm. Time resolution of 300s (5 min). In each simulation, a 1-phase 100 kW PV system was placed at each of the node of buses 64, 135 and 83. S matrix size is 229x9.

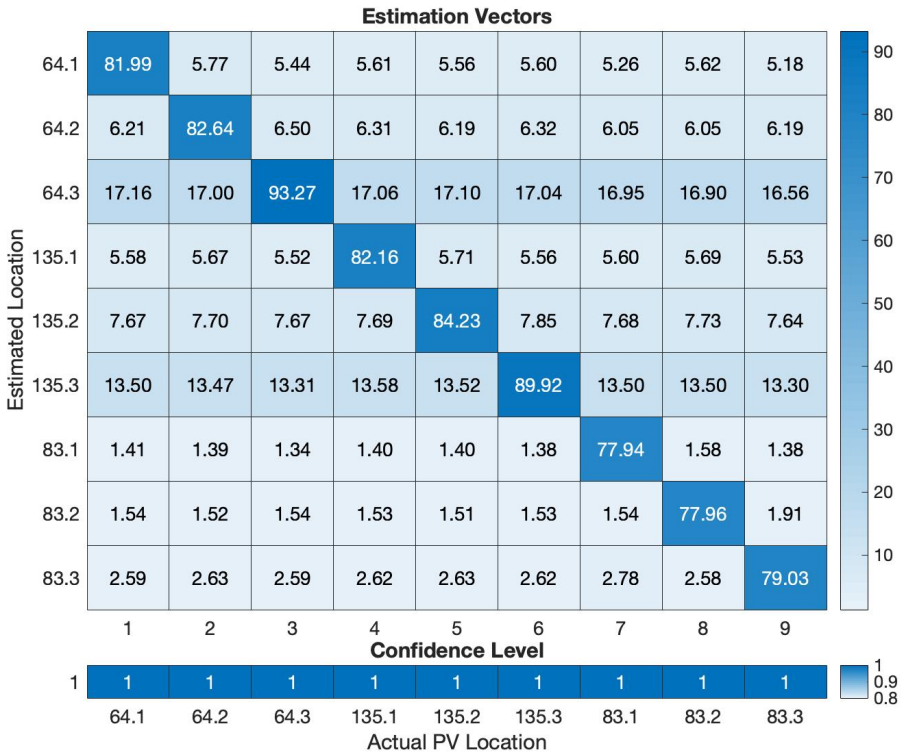


Figure 46 – Estimation results for 9 different simulations within the time window 12:25 to 12:30pm. Time resolution of 300s (5 min). In each simulation, a 1-phase 100 kW PV system was placed at each of the node of buses 64, 135 and 83. S matrix size is 18x9.

Some biases towards certain nodes appear when measures are excluded – in this case towards node 64C – but the estimation can still be accurate. This is in fact the expected result, as we are eliminating essential characteristics of the sensitivity vectors. We can then conclude that the method is robust even while having a significant portion of voltage data missing.

In the particular cases in which the resulting $\boldsymbol{S}$ matrix is not full rank due to the elimination of row, we need regularization techniques to overcome the singularity issue. In this research, we use the Tikhonov regularization. This approach consists of slightly modifying the least-squares solution by considering a penalization matrix $\boldsymbol{\Gamma}$. This matrix has the same size as $\boldsymbol{S}$, and it is characterized by having all components equal to zero except those of the columns associated to one of the sensitivity vectors that is linearly dependent from others after eliminating a row. That is, we are penalizing the estimation of that sensitivity vector to allow the model to compute the inverse. The estimation model using the Tikhonov regularization becomes:

$$\hat{\boldsymbol{x}} = (\boldsymbol{S}^T\boldsymbol{S} + \boldsymbol{\Gamma}^T\boldsymbol{\Gamma})^{-1}\boldsymbol{S}^T\boldsymbol{d} \tag{5.2}$$

such that $\boldsymbol{S}^T\boldsymbol{S} + \boldsymbol{\Gamma}^T\boldsymbol{\Gamma}$ remains full rank. Let us take as example when we are missing the voltage data of bus 634B and we are considering node 633B as a potential PV location candidate. In this case, the resulting sensitivity vectors are proportional to each other. The regularization let us perform the estimation and still get accurate results. Figure 47 shows the impact of this technique on the mentioned example.

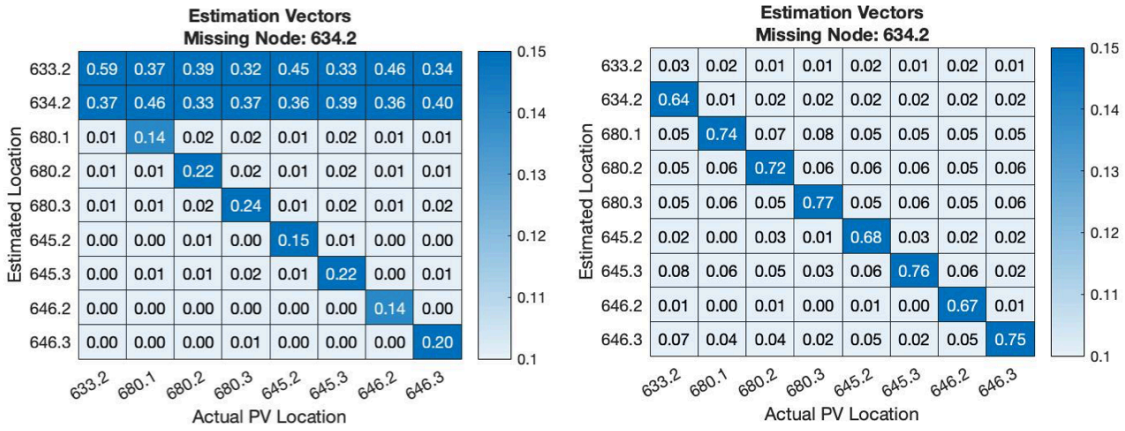


Figure 47 – Estimation results before (left) and after (right) applying Tikhonov regularization.

The method is now able to predict correctly all the PV locations. Note that the penalization has been applied to location 633B, so that any PV located at either 633B or 634B will be predicted as 634B. The algorithm will estimate the node that is not penalized by $\boldsymbol{\Gamma}$ in case of a PV at either of both locations. In other words, with the elimination of one row, we only can predict L-1 locations of PV: we either predict location 633B or 634B, but not both at the same time. Thus, under these circumstances we can only predict clusters of nodes – regions in the feeder – where the PV is most likely located.

5.3 Algorithm Limitations

This section encapsulates all the limitations of the method that have been discussed along this study, both based on the method assumptions and numerical limitations:

a) The method assumes a perfect model of the distribution feeder and no noise in the measurements. This might not be the case for utilities models and errors in the sensitivities and sensors may impact the estimations. The impact of errors in sensitivities is similar to the noise in measurement, as the methods provides the solution of components that best align measures with sensitivity vectors. This study shows there is a trade-off between states and the acceptable ratio of errors.

b) For certain combinations – those in conflict with VRE – of multiple PV with high-power injections that exceeds the expected impact of VRE on the estimation, the accuracy of the method will be compromised as it reaches its limits of normal operation conditions.

c) The method takes as input the time-point voltage samples. However, some measurement devices only provide the average voltage over a sampling period, which leads to a different measurement vector that can impact the accuracy of estimations.

d) The method works best with large punctual PV injections that need to be selected manually from the voltage profile. For some points in time that there are no voltage changes due to PV, the results lead to inaccurate estimation. A clear sunny day is a good example of this situation.

e) This study does not include the impacts of continuous voltage regulating devices like advanced inverters (e.g. with volt-var), which may be very significant in terms of sensitivities and estimation results.

CHAPTER 6. CONCLUSIONS

6.1 Conclusion

A method has been described for the estimation of the PV locations under the presence of voltage regulation equipment, based on a limited number of available voltage magnitude measurement streams while preserving high accuracy rates. The high estimation confidence level corroborates the results of the analysis. The estimation model enables recovering the most likely location of multiple PV systems given an accurate distribution system model and power sensitivities, in addition to the voltage magnitude measurements that utilities can monitor regularly. The limitations of an imperfect distribution model and noise measurement have been studied, as well as the multiple factors that impact the method's accuracy.

The impact of capacitors and voltage regulators has been explored. In particular, the effects on the estimation algorithm has been analyzed. The knowledge of the expected impact is used to determined when the VRE acted. Depending on how the structure of VRE is in terms of the PV sensitivities, we may use one of the two methodologies presented in this study. If the increase of voltage is linearly independent from the expected voltage increment due to PV, we may proceed with the incorporation of the VRE sensitivity as an

addition into the $\boldsymbol{S}$ matrix. On the contrary, if the impact is linearly dependent, it is required to perform a second estimation to determine the voltage increase due to VRE and subtract it from the measurement vector to obtain an accurate estimation. The simulations carried out on the IEEE 13-bus and IEEE 123-bus test feeders corroborate the theory presented and provided highly accurate results in terms of PV injection estimation, and VRE actions detection.

6.2 Future Work

Future work will address more realistic situations of the PV estimation, such as enabling the method to estimate under averaged voltage magnitude measurements provided by voltage sensors. The method requires to manually select points in time with a PV injection to perform the estimation, so an algorithm to preselect these points in a timeless manner needs to be developed to automatize the PV location prediction process. Finally, the methodology presented does not consider the impacts of continuous voltage regulating devices like advanced inverters (e.g. with volt-var). The effect of these regulating devices will be significant on the estimation, so an additional framework that takes into account continuous VRE needs to be developed.

REFERENCES

[1] L. Blakely, M. J. Reno, and J. Peppanen, “Identifying Common Errors in Distribution System Models,” IEEE Photovoltaic Specialists Conference (PVSC). 2019

[2] X. Zhang, S. Grijalva, “A Data-Driven Approach for Detection and Estimation of Residential PV Installations. IEEE Trans Smart Grid, 2016;7:2477–85.

https://doi.org/10.1109/TSG.2016.2555906.

[3] M.U. Qureshi, S. Grijalva, M.J. Reno, J. Deboever, X. Zhang and R.J. Broderick, “A Fast, Scalable Quasi-Static Time Series Analysis Method for PV Impact Studies using Linear Sensitivity Model”, IEEE Transactions on Sustainable Energy, Vol 10, No. 1 Jan, 2019.

[4] A. Kumar, S. Grijalva, J. Deboever, J. Peppanen, M. Rylander, “Mathematical Representation of Voltage Regulation Impact on Distribution Feeder Voltages, 46th IEEE Photovoltaics Specialists Conference (PVSC), Chicago, IL, June 16-21. 2019.

[5] Reinaldo Tonkoski, IEEE Student Member, and Luiz A. C. Lopes, IEEE Senior Member. “Voltage Regulation in Radial Distribution Feeders with High Penetration of Photovoltaic”. 2008.

[6] K. Mason, M.J. Reno, S. Vejdan, S. Grijalva, “A Deep Neural Network Approach for Behind-the-meter Residential PV Size, Tilt and Azimuth Estimation”, Solar Energy 2020; 196:260–9. https://doi.org/10.1016/J.SOLENER.2019.11.100

[7] D.L. Donaldson, D. Jayaweera, “Effective solar prosumer identification using net smart meter data”, Elsevier Journal of Electrical Power and Energy Systems, Vol 118, 2020.

[8] J. M. Malof, K, Bradbury, L. M. Collins, R. G. Newell, “Automatic Detection of Solar Photovoltaic Arrays in High Resolution Arial Imagery”, http://dx.doi.org/10.1016/j.apenergy.2016.08.191

[9] K, Bradbury, R. Saboo, J. Malof, T. Johnson, A. Devarajan, W. Zhang, “Distributed solar photovoltaic array location and extent data set for remote sensing object identification”. Figshare; 2016. http://dx.doi.org/10.6084/ m9.figshare.3385780.v1

[10] F. Wang, K Li, X Wang and al “A Distributed PV System Capacity Estimation Approach Based on Support Vector Machine With Customer Net Load Curve Features”, Energies 2018, 11, 1750. https://dx.doi.org/10.3390/en11071750

[11] J. Tan, C. Deng, W. Yang, N. Liang, F. Li, “Ultra-Short-Term Photovoltaic Forecasting in Microgrid based on Adaboost Clustering”. 2017 http://dx.doi.org/10.7500/AEPS20170217006

[12] S. Pouraltafi-Kheljan, M. Göl, “Power Generation Nowcasting of the Behind-Meter Photovoltaic Systems”, arXiv.2001.02157v1

[13] X. Shi, J. Yao, Y. Ye, H. Feng, X. Hu, “Estimation of Invisible Distributed PV Power Generation From Bus Load”, iSPEC 2019. https://dx.doi.org/10.1109/iSPEC48194.2019.8974968

[14] M. Hernandez, J. Peppanen, J. Deboever, M. Rylander, “Enhanced Load Modeling: Survey Results on Industry Load Modeling Practices”

[15] F. Bu, K. Dehgahnpour, Y. Yuan, Z. Wang, Y. Guo, “Disaggregating Customer-level Behind-the-Meter PV Generation Using Smart Meter Data”, arXiv.org, Electrical Engineering and Systems Science, Signal Processing, arXiv:2009.00734.

[16] F. Kabir, N. Yu, W. Yao, R. Yang, Y. Zhang, “Joint Estimation of Behind-the-Meter Solar Generation in a Community,”, 2020 IEEE Transactions on Sustainable Energy.

[17] S. Grijalva, A. U. Khan, J. S. Mbeleg, and C. Gomez-Peces, M. J. Reno, L. Blakely, “Estimation of PV Location in Distribution Systems based on Voltage Sensitivities”, 2020 North American Power Symposium.

[18] M.U. Qureshi, S. Grijalva, M.J. Reno, “A Fast Quasi-Static Time Series Simulation Method for PV Smart Inverters with VAR Control using Linear Sensitivity Model”, 7th World Conference on Photovoltaic Energy Conversion.

[19] M. Shiroie, S. H. Hosseini, “Observability and Estimation of Transformer Tap Setting with Minimal PMU Placement”, 2008 IEEE Power and Energy Society General Meeting.

[20] E. Handschin, E. Kliokys “Transformer Tap Position Estimation and Bad Data Detection using Dynamic Signal Modelling,” 1995 Transactions on Power Systems, Vol 10, No. 2, May 1995.

[21] Sengupta, Manajit; Andreas, Afshin (2014): Oahu Solar Measurement Grid (1-Year Archive): 1-Second Solar Irradiance; Oahu, Hawaii (Data). National Renewable Energy Laboratory. https://data.nrel.gov/submissions/11.

[24] R. C. Dugan, A. Ballanti, “The Open Distribution System Simulator (OpenDSS),” Electric Power Research Institute, Inc. March 2016

[25] M. J. Reno, K. Coogan, “Grid Integrated Distributed PV (GridPV) Version 2,” Sandia National Laboratories, November 2014

[26] M. Brown, M. Biswal, S. Brahma, S. J. Ranade, H. Cao, “Characterizing and Quantifying Noise in PMU data,” 2016 IEEE Power and Energy Society General Meeting.